An Australian approach to setting air quality standards:

Consultation draft

As at September 2009

NOTE

This document has been developed by a project team of government officers from health and environment portfolios, at the request of the National Environment Protection Council (NEPC). The paper is provided as a basis for discussion and does not necessarily reflect the views of the NEPC Committee. The paper does not carry the endorsement of the NEPC nor any member government.

The contents of the paper should not be used for any purpose other than as a basis for providing comment to the NEPC Committee.

An Australian approach to setting air quality standards: Consultation draft

© Copyright vests in the Commonwealth of Australia and each Australian State and Territory.

10 digit ISBN 1–921173-60-2 (Print)
13 digit ISBN 978-1-921173-60-8 (Print)

10 digit ISBN 1–921173-59-9 (Electronic)
13 digit ISBN 978-1-921173-59-2 (Electronic)

This work is copyright. It has been produced by the National Environment Protection Council (NEPC). Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced by any process without prior permission from the NEPC Service Corporation. Requests and enquiries concerning reproduction and rights should be addressed to the Executive Officer, NEPC Service Corporation, Level 5, 81 Flinders Street, Adelaide SA 5000.

CONTENTS

FIGURES

TABLES

1 INTRODUCTION

1.1 NATIONAL ENVIRONMENT PROTECTION COUNCIL

The National Environment Protection Council (NEPC) and its operations are established by the *National Environment Protection Council Act 1994* (Commonwealth) and corresponding state and territory Acts. The NEPC is a national body empowered by the NEPC Act to develop and make National Environment Protection Measures (NEPMs). The NEPC is a statutory entity within the Environment Protection and Heritage Council (EPHC), whose role is to harmonise environmental protection approaches across Australia. The NEPC and EPHC comprise ministers representing each of the participating governments.

1.2 NATIONAL ENVIRONMENT PROTECTION MEASURES

A NEPM is a legislative instrument designed to protect particular aspects of the environment. It may have one or more goals, standards and protocols; it may also contain guidelines.

The NEPC Act requires the NEPC to have regard to the Intergovernmental Agreement on the Environment (IGAE) 1992 when making NEPMs. The IGEA is published as a Schedule to the NEPC Act (1994) establishes the general provisions for the making of NEPMs. As stated in the IGAE, the objectives of NEPC are to ensure that:

- people enjoy the benefit of equivalent protection from air, water and soil pollution, and from noise, wherever they live
- decisions by businesses are not distorted and markets not fragmented by variations between jurisdictions in relation to the adoption or implementation of major environment protection measures.

The IGAE also requires the development of NEPMs to take into account the precautionary principle.

As defined by the NEPC Act, a national environment protection standard 'consists of quantifiable characteristics of the environment against which environmental quality can be assessed'. In the past, standards have been numerical values that have taken a variety of forms; these include:

- standards with a compliance goal, and with specified monitoring and reporting protocols
- advisory reporting standards
- monitoring investigation levels.

1.2.1 Ambient Air Quality NEPM

In 1998, the NEPC made the Ambient Air Quality (AAQ) NEPM. This NEPM establishes a nationally consistent framework for the monitoring and reporting of air quality, and establishes national ambient air quality standards that apply in all states and territories and land controlled by the Commonwealth. These standards were the first step in developing a more consistent national approach to air quality management. The NEPM was varied in 2003 to incorporate advisory reporting standards for $PM_{2.5}$ (i.e. particulate matter of 2.5 μm or less).

The AAQ NEPM applies to what are known as the criteria pollutants. These are common pollutants that include particles (as PM_{10} and $PM_{2.5}$), ozone, carbon monoxide, nitrogen dioxide, sulfur dioxide and lead; they arise from multiple sources and are widespread in the ambient air environment. The criteria pollutants have been monitored by environment agencies on a routine basis for many years, and there are large databases of monitoring data. The health effects arising from exposure to these pollutants are well documented. Although associated with small individual risks, the health effects translate to large population risks and public health burden because the

entire population is exposed. Similar health effects are observed for most of the pollutants, and some groups are more susceptible to their effects.

With the exception of the advisory reporting standard for $PM_{2.5}$, the standards in the AAQ NEPM are compliance standards. The level of protection incorporated in advisory reporting standards is the same as that for a compliance standard, which means that the assumptions and overall approach to the standard setting process are also the same. However, advisory reporting standards do not have an associated goal that sets a timeframe for compliance. The monitoring protocol associated with an advisory reporting standard establishes a reference method, and requirements for monitoring and reporting, but gives jurisdictions flexibility about the timing and extent of monitoring they conduct. Any data collected can be assessed against the advisory reporting standard, which provides a benchmark against which the risk of adverse health effects arising from exposure to $PM_{2.5}$ can be assessed.

1.2.2 Air Toxics NEPM

In December 2004, NEPC made the Air Toxics NEPM. This NEPM establishes a nationally consistent framework for monitoring and reporting on air toxics in Australia.

Air toxics are a large group of substances that have been associated with various health effects (including cancer, respiratory irritation, developmental and reproductive effects). They are hazardous pollutants that arise from a diverse range of sources including industrial processes, motor vehicles and domestic sources. Concentrations of air toxics in ambient air are generally low, but are highly variable across an airshed compared with concentrations of criteria pollutants, although their relative toxicity can be much greater. In general, elevated levels of air toxics will only be observed near their sources, and such locations are often referred to as peak sites or 'hotspots'.

Certain groups within the population are more exposed to air toxics than the general population. Historically environment agencies have not routinely monitored air toxics, and monitoring data are limited. Health data on air toxics, including unit risk factors, are derived largely from occupational or animal studies, rather than from studies on the general population.

When the Air Toxics NEPM was made, there were insufficient data to set either compliance standards or advisory reporting standards for the pollutants covered by the NEPM. Monitoring investigation levels were developed that set numerical values protective of human health (at upper-bound risk levels). If these values are exceeded over the appropriate averaging period, then investigation may be triggered — this can involve further monitoring and assessment of the circumstances that may have led to the levels being exceeded. Measurements are made at locations where significantly elevated ambient levels of the pollutant might be expected, and where significant numbers of people might be exposed.

The level of health protection associated with the investigation levels differs from that associated with compliance standards and advisory reporting standards. The monitoring investigation levels have been established as trigger levels for further investigation; they reflect an upper bound of risk that should not be exceeded. In contrast, the compliance and advisory reporting standards for the criteria pollutants in the Ambient Air Quality NEPM incorporate a significant margin of safety that should protect most of the population, including sensitive groups, from the adverse effects of exposure to those air pollutants. Given that many of the criteria pollutants do not have a threshold for effect, there is likely to be some degree of residual risk associated with any standard.

1.3 A PARTNERSHIP APPROACH BETWEEN ENVIRONMENT AND HEALTH

In the past, the NEPM standard setting processes in Australia used a variety of frameworks and methods, each with its own merits and disadvantages. Following the successful work undertaken by the Risk Assessment Task Force (RATF) and the Risk Assessment Working Group (RAWG) and

subsequently in the application of risk assessment to the development of standards for $PM_{2.5}$., the absence of an overall agreed methodology was reflected in the considerable debate across the health and environment sectors, and with other key stakeholders during NEPCs standard-setting process.

The areas of debate go beyond the broad standard-setting framework. They involve technical issues such as:

- margins of safety required in the standards that are set to ensure protection of sensitive groups
- the relative use of both epidemiological and toxicological data in the hazard assessment
- the approaches to dealing with non-threshold pollutants, such as particles, ozone and carcinogens.

The main aim of setting NEPM standards is to prevent adverse health impacts from air pollution, and to provide adequate protection for all Australians. Achieving this aim requires a comprehensive knowledge of the health effects of air pollutants. Standard setting in Australia — in particular, the assessment of potential health hazards associated with exposure to air pollution — requires collaboration mainly between the environment and the health sector. Such collaboration has been in place for as long as air quality standards have been set in Australia, but the approach has not always been consistent.

1.4 THE STANDARD SETTING WORKING GROUP

The Standard Setting Working Group (SSWG), which operates under the auspices of the EPHC Air Quality Working Group, was asked to develop an overall framework and methodology for setting air quality standards in the NEPC context. Health experts worked with experts from the environment sector to develop this methodology, with the SSWG being co-chaired by representatives from both sectors. Members were nominated for their demonstrated expertise and experience in, or understanding of:

- epidemiology or toxicology, and in the application of these studies in setting standards
- risk assessment, especially in relation to setting air quality standards
- policy underpinning standard setting
- air quality data.

1.4.1 Scope of the work

The objectives of the SSWG are to scope and assess approaches to the derivation of air quality standards in Australia, and to make recommendations on an agreed methodology to the NEPC.

The NEPC requested that the overall standard-setting framework be based on a risk assessment approach. The proposed methodology aims to reach agreement between the health and environment sectors, and consultation with key stakeholders on:

- the level of health protection to be built into standards
- the application of uncertainty or safety factors
- approaches to dealing with non-threshold pollutants
- approaches to exposure assessment
- equity and social justice issues
- application and approaches to cost–benefit analysis.

Although the potential for a mixture of air pollutants to impact on health is acknowledged, this document does not provide guidance on how to address potential combined effects from air pollutants in a standard setting process. This is consistent with the current approaches used internationally, where air quality standards are established for individual pollutants.

1.4.2 Basis for the methodology

Recommendations from a number of Australian documents were considered when formulating the recommended framework, including:

- *Report of the Risk Assessment Task Force — Prepared for the National Environment Protection Council* (RATF 2000). The taskforce investigated the potential for the use of health risk assessment in the NEPC context, particularly in relation to the review and establishment of ambient air quality standards for criteria pollutants addressed in the AAQ NEPM.
- *Report of the Risk Assessment Working Group — Prepared for the National Environment Protection Council* (RAWG 2002). Following receipt of the RATF report, the NEPC requested that a report on the 'future actions' arising from the report be prepared. The RAWG report is a significant supplement to the RATF report; it is based on a review of overseas risk assessment approaches and an analysis of air monitoring information provided by jurisdictions.
- *Ambient Air Quality Standards Setting – An Approach To Health-Based Hazard Assessment* (NHMRC 2006): In 2006 the National Health and Medical Research Council (NHMRC) — in conjunction with the Environmental Health Committee (enHealth) and the Australian Health Protection Committee — conducted a national expert workshop. The workshop was convened to consider in depth the issues of health-hazard assessment for setting health-based air quality standards, and to arrive at a methodology supported by both the health and the environment sector.

In addition to proposed Australian methodologies, the SSWG also considered standard-setting procedures followed by the United States Environmental Protection Agency (US EPA), the United Kingdom Department for Environment, Food and Rural Affairs (DEFRA), the European Union, the California Environmental Protection Agency (Cal EPA) and its Office of Environmental Health Hazard Assessment (OEHHA), and the World Health Organization (WHO).

1.4.3 Target audience and purpose

This document has been written to help those working in health and the environment to develop ambient air quality standards, by providing a clearly defined framework and set of recommendations. It is expected to be particularly relevant to professionals developing NEPM standards.

This document provides a clearly defined framework for the standards setting process; however, it is important that the document is not seen as a recipe book or a set formula that can be simply followed. The standard setting process is complex and pollutant-specific, that is, a different process may be required for each pollutant. There is often the need for expert judgement to be employed, and it is important that decisions are transparent and well documented.

The purpose of this consultation version of the technical document is to encourage discussion on the issues, processes and recommendations put forward here, with the aim of developing an agreed framework and methodology for the purpose of setting air quality standards. The feedback provided will help to ensure that the process and its outcomes are as transparent as possible. The paper discusses approaches to health hazard assessment, exposure assessment, risk characterisation and science policy issues in setting air quality standards in the NEPC context.

1.4.4 Structure of this paper

Section 2 of this paper provides the background to setting air quality standards and frameworks that have been used to derive them. This section also provides a diagrammatic presentation of the recommended framework (see Figure 2.3) on which the methodology follows.

Sections 3–5 of this paper discuss the technical issues involved in the risk assessment components of this standard setting framework, and the international and previous Australian responses to these issues. These sections also recommend an approach for use in the NEPC context.

Section 6 of this paper summarises the pertinent points from the risk assessment process, described in sections 3-5, which lead to the formation of the range of potential risk based guidelines.

Section 7 discusses a range of policy issues included in the NEPC Act that guide the scope of the risk assessment process and that need to be considered in setting air quality standards and recommends guidance on addressing these considerations.

The NEPC Act requires social, environmental and economic factors to be considered in setting standards for air quality. These considerations inform the decision-making process as part of the risk management component of standard setting.

Section 8 provides information on how comments or submissions on this document can be provided. The submissions will assist in finalising the document for EPHC consideration.

2 FRAMEWORK FOR SETTING AIR QUALITY STANDARDS

2.1 AIR QUALITY STANDARDS

Developing a NEPM standard involves taking into account health, environmental, social and economic considerations. Many international air quality guidelines are based purely on health considerations. Standards usually have a compliance component that sets a timeframe for meeting the standard; this component is often legally binding. Internationally, air quality standards and guidelines are usually developed through a risk assessment process that is either qualitative or quantitative, whereas in Australia, approaches have differed between NEPMs, as discussed below.

Some definitions

The main aim of setting NEPM standards is to prevent adverse health impacts from air pollution, and to provide adequate protection for all Australians. Hence, standard setting for air pollutants requires collaboration between the environment and the health sectors. This collaboration enabled the National Health and Medical Research Council (NHMRC) to develop and publish in 2006, a document entitled *Ambient Air Quality Standards Setting – An Approach To Health-Based Hazard Assessment* which outlines the process for developing **health-based air quality guideline values**. These values can be used to derive standards for the AAQ and Air Toxics NEPM by the application of several important policy considerations including community consultation.

The standards in the AAQ NEPM are **compliance standards** with an associated goal that sets a timeframe for compliance. The exception is the **advisory reporting standard** for $PM_{2.5}$, which has the same level of protection as that for a compliance standard, but gives jurisdictions flexibility about the timing and extent of monitoring they conduct.

When the Air Toxics NEPM was made, there were insufficient data to set either compliance standards or advisory reporting standards for the pollutants covered by the NEPM. **Monitoring investigation levels** protective of human health were developed instead. If these values are exceeded over the appropriate averaging period, then investigation may be triggered.

The level of health protection associated with the investigation levels differs from that associated with compliance standards and advisory reporting standards. The monitoring investigation levels have been established as trigger levels for further investigation; they reflect an upper bound of risk that should not be exceeded. In contrast, the compliance and advisory reporting standards for the criteria pollutants in the Ambient Air Quality NEPM incorporate a significant margin of safety that should protect most of the population, including sensitive groups, from the adverse effects of exposure to those air pollutants. Given that many of the criteria pollutants do not have a threshold for effect, there is likely to be some degree of residual risk associated with any standard. The NEPM standards are meant to apply to ambient air only and may not be appropriate for use as indoor air quality standards.

2.2 OVERVIEW OF INTERNATIONAL FRAMEWORKS FOR SETTING AIR QUALITY STANDARDS

Many countries develop air quality standards and guidelines. The standards are used in different legislative frameworks, and the approaches to developing the standards or guidelines can differ.

In the United States, the United States Environmental Protection Agency (US EPA) establishes National Ambient Air Quality Standards (NAAQS) for the criteria pollutants, and reference concentrations (RfCs) for air toxics. The NAAQS are legally binding on states, which must develop State Implementation Plans to meet the NAAQS. The RfCs are guideline values. These NAAQS and RfCs are based solely on the consideration of health effects; economic considerations are not explicitly taken into account.

The World Health Organization (WHO) bases its air quality guidelines solely on health considerations. The WHO guidelines are used as the basis of air quality standards in many countries; for example, they are used by the European Union as the basis of limit values. However, WHO recommends that social and economic issues for each country or region be considered in setting standards.

Both the US EPA and WHO assess both epidemiological and toxicological data. For the criteria pollutants, the focus in recent years has been on the use of population-based epidemiological studies as the basis for setting air quality standards. The results of controlled human exposure studies and of toxicological studies add to the weight of evidence for an adverse effect linked to exposure to the pollutant under consideration; they also provide evidence for biological plausibility of the effects observed in the epidemiological studies. Strict criteria have been developed to guide the selection and evaluation of studies, as part of systematic review of the literature. Internationally, a weight-of-evidence approach is used consistently.

The WHO and the United Kingdom use expert panels to review health information while the US EPA staff work with external experts to conduct systematic reviews of the literature, which are then reviewed by an independent panel of experts. The use of expert panels ensures that the relevant expertise is used in developing standards; it also provides independent review of the information on which standards are based.

2.2.1 Use of risk assessment

In the United States, NAAQS for criteria air pollutants are developed using quantitative risk assessments and city-specific data (for a number of cities across the country). A semi-quantitative approach is used in developing RfCs. Risk assessments are undertaken for each of these cities individually using local data.

To assess achievability of guideline values, WHO conducts a qualitative assessment of the health data and reviews air quality data from various parts of the world. Similarly, the United Kingdom expert panel conducts qualitative assessments of the literature, and evaluates air quality in the United Kingdom, to assess whether air quality standards are achievable.

2.3 PREVIOUS APPROACHES TO SETTING AIR QUALITY STANDARDS IN AUSTRALIA

2.3.1 Ambient Air Quality NEPM

A qualitative risk assessment process was used to develop the standards for the AAQ NEPM. An independent review was commissioned to asses the state of knowledge of the human health effects of the six pollutants covered by the NEPM, and to identify adverse health impacts on both the general population and on any susceptible subgroups. A technical review panel — consisting of government, industry and community health and medical experts — was established to review the consultant's reports and make recommendations on possible standards. The report was internationally peer reviewed, and its conclusions were supported.

The outcome of the consultancy was a series of recommendations on the ambient levels (or pollution concentration ranges) that would provide protection from the adverse effects of each pollutant on susceptible subgroups in the population.

The main inputs to the assessment process were:

- the outcomes of the health review
- an exposure assessment
- an examination of the air quality management or 'control' options and their associated costs for achieving the proposed standards

- consideration of the benefits, typically in terms of avoided health costs, associated with each of the standards.

As a general principle, uncertainty (or safety) factors were not used in developing AAQ NEPM standards, except where there was uncertainty about the existence of a health effects threshold. Such factors are often used in setting standards, to account for uncertainty in the data.

2.3.2 $PM_{2.5}$ variation

In 2001, the NEPC commenced a review on the need for $PM_{2.5}$ standards. The review included information from Australian and international studies on adverse health effects and exposure-response data. On the basis of the review, the NEPC recommended that a $PM_{2.5}$ standard be included in the AAQ NEPM, and that a variation to the NEPM be undertaken. The standards for $PM_{2.5}$ were developed through a full quantitative risk assessment process that used air quality data from Brisbane, Melbourne, Perth and Sydney. The risk assessment process followed the process recommended by the NEPC RATF (NEPC 2001).

Australian and international literature on the health effects attributable to $PM_{2.5}$ was systematically reviewed. Sensitive groups and exposure–response functions were identified. The health data were derived largely from epidemiological studies supplemented by laboratory studies on individual or groups of subjects, mainly from the United States and Europe. The exposure–response functions for a number of health outcomes and sensitive groups were identified for use in the risk characterisation. In conducting the exposure assessment for the development of the standards for $PM_{2.5}$, the approach taken by the US EPA in their standard-setting process was followed. This mirrored the approach used to assess exposure in the epidemiological studies from which the health effects and dose–response data were derived. The outcome of the risk assessment process provided estimates of the number of health outcomes avoided if $PM_{2.5}$ levels could be reduced to meet each of the potential standards. The costs and benefits associated with each case were assessed to develop a recommendation on the standard. The costs of hospital admissions used in the cost–benefit analysis were obtained from the Australian Institute of Health and Welfare (AIHW). No attempt was made to provide costing on the mortality estimates because at the time Australia had no agreed method for doing this. The risk assessment used as the basis for the $PM_{2.5}$ standards was internationally peer reviewed.

2.3.3 Air Toxics NEPM

Given the lack of data for air toxics in Australia when the Air Toxics NEPM was made, it was not possible to use a quantitative risk assessment process to develop standards for the air toxics under consideration. Instead, existing overseas standards and guidelines were assessed for their applicability as investigation levels — that is, levels that, if exceeded, may trigger some form of further investigation — in the Australian context. The only overseas standards or guidelines selected were those that had been derived for a similar purpose to that proposed in the Air Toxics NEPM. Various criteria were applied in selecting overseas standards or guidelines; these criteria were consistent with the approaches taken overseas in the evaluation of key studies and standards.

As part of the development of the proposed NEPM, the health effects of the five air toxics under consideration were extensively reviewed. These reviews, which were appraised by representatives of the health sector in Australia, highlighted the key epidemiological and toxicological studies used as the basis for overseas standards. Consultation with the health sector indicated that the identified health endpoints are seen as appropriate as the basis for air quality standards for air toxics in Australia. The endpoints were also consistent with those used overseas in the derivation of standards and guidelines for ambient air quality.

2.3.4 Air Toxics Tier 2

In determining the scope of the Air Toxics NEPM, a phased approach was proposed, whereby hazardous pollutants not included in the original NEPM could be considered for incorporation at a later date (referred to as 'Tier 2'). When making the Air Toxics NEPM, the NEPC also established a working group to develop a methodology to prioritise pollutants; the methodology is available on the EPHC website.[1] The main task of the working group was to develop and trial a method to screen and prioritise pollutants, the aim being to identify priority pollutants in ambient air that are of national significance in terms of exposure, and that may pose a risk to human health. The emphasis in determining the health risk is on public health rather than occupational health. The priority pollutants identified may then be considered for incorporation in the NEPM, and standards or guidelines may be developed for them.

The working group reviewed international approaches to prioritising air pollutants and then developed a risk-based method with two components:

- *hazard identification* — which considers cancer classification and potency, respiratory effects, reproductive and developmental effects, other non-cancer effects, and emerging issues. Each pollutant is individually scored on hazard on this basis; as outlined in the review, these questions are consistent with those used in other countries for this purpose
- *exposure estimation* — which involves adopting surrogate measures of population exposure to air toxics; these measures rely on air emissions inventories (in particular the National Pollutant Inventory) for each of the air toxics and on atmospheric-fate data; they are consistent with those used in international schemes to prioritise air pollutants. Individual scores are applied.

Pollutants are first scored and ranked separately against the two components, which are then combined to give the final relative-risk ranking.

2.3.5 Common features

In Australia, a systematic approach has been used to evaluate the literature, both overseas and in previous standard-setting processes. The approach involves establishing strict guidelines to guide the selection of studies for inclusion in the review of the literature. A weight-of-evidence approach is used. For the criteria pollutants, the results of epidemiological studies play a central role in determining the risk these pollutants pose to human health.

Risk assessment is used by all international agencies, and has been used successfully in Australia in developing standards for $PM_{2.5}$. The approach used in Australia is consistent with the quantitative approach used by the US EPA, and allows an assessment of costs and benefits associated with any potential standards. This is required under the NEPC Act for the development of any air quality standards in Australia.

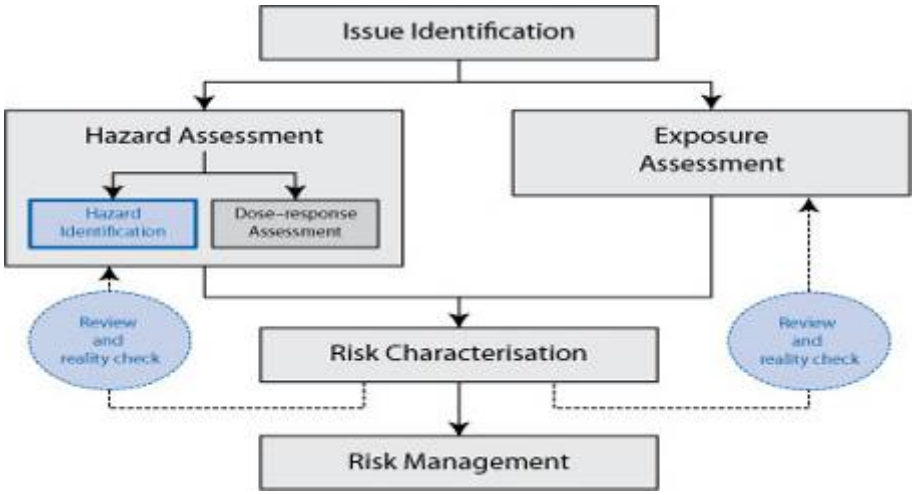
2.4 AUSTRALIAN APPROACHES TO RISK ASSESSMENT

Under the NEPC Act, standard setting involves a range of considerations — environmental, health, technical, social, economic, political, legislative and cultural. Standards are primarily based on the protection of human health or environment.

Risk assessment has been used widely in Australia to evaluate risk posed to human health or the environment arising from environmental hazards (e.g. air pollution) for many years. The RATF developed a risk assessment framework specifically for use in the development of air quality standards. This framework is consistent with that developed by enHealth in 2002 for assessing the impact of environmental hazards on human heath (shown in Figure 2.1, below), and with risk assessment approaches used worldwide.

[1] *http://www.ephc.gov.au*

Figure 2.1 EnHealth environmental health risk assessment framework; enHealth 2002



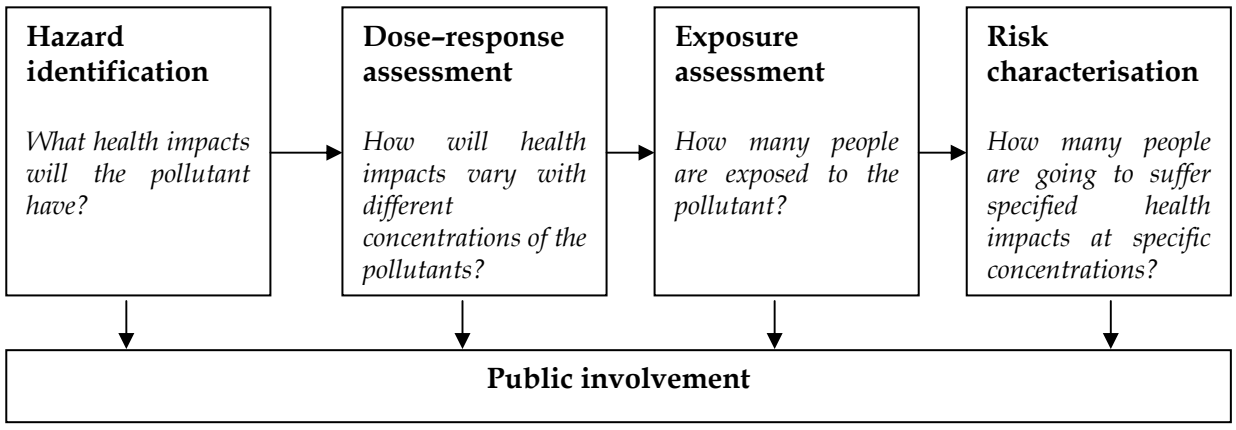
It is important to separate the technical and scientific risk assessment process set out in Figure 2.1 from the risk management process. The risk management process takes into account a range of factors (including social and economic factors) that are not considered in the risk assessment process.

Figure 2.2 shows the systematic four-stage risk assessment framework recommended by the RATF for setting air quality standards in Australia. The framework is applicable to a wide range of pollutants and health effects. One key component is the importance of public involvement in the risk assessment process, to ensure that the process is transparent, and that its outcomes are understood and accepted.

The stages of dose–response assessment, exposure assessment and risk characterisation may incorporate mathematical models. The specific models used for dose–response assessment and exposure assessment depend on the pollutant (and its biochemical and physical properties) and the health endpoint chosen.

The outputs from risk assessment are used in the risk management phase (as set out in Figure 2.1). In the risk management phase, regulators consider the results of the risk assessment stages outlined above, apply the precautionary principle and take into account social and economic factors.

Figure 2.2 Four-stage framework of health risk assessment



As shown in Figure 2.2, risk assessment is often considered to have five stages, including an initial Identification stage. The RATF specifically separated issue identification from the health risk assessment framework in the standards development process, because it is the necessary first step in that process. A standard-setting process will only be initiated by the NEPC if an 'issue' with that pollutant has previously been identified. If there is no disagreement about the need for a standard or about the level at which a standard should be set, then there is no need for an issue identification step as the first stage in the health risk assessment.

NHMRC has recommended that the enHealth risk assessment model be used for setting national ambient air quality standards (NHMRC 2006). This approach to risk assessment is consistent with the frameworks used overseas; it has been used in the development of air quality standards in Australia in the past for $PM_{2.5}$; it has also been used internationally by WHO and US EPA.

In reviewing standard setting methodologies the SSWG developed a framework for the NEPC standard setting process. The framework is shown in Figure 2.3 and is consistent with international and previous Australian approaches to standard setting, including those of the RATF and enHealth. The shaded boxes depict the main stages discussed in this document.

The stage that identifies and prioritises pollutants for which standards are required is not discussed in this document. It is assumed that the AAQ NEPM review will identify any criteria pollutants that require revised standards. The NEPC has developed a method for prioritising additional air toxic pollutants, besides the five already identified, for possible future inclusion in the Air Toxics NEPM.

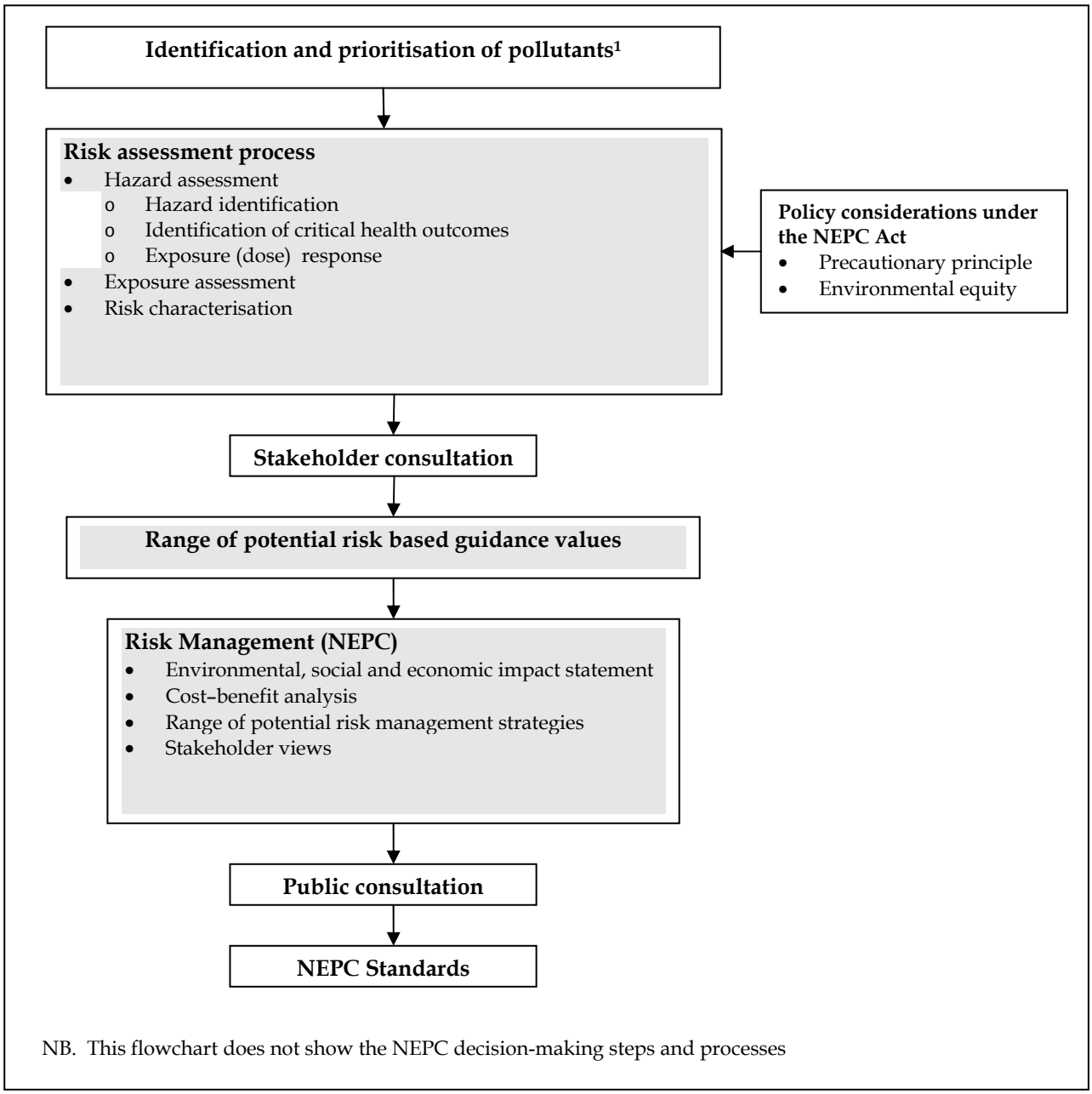
In Figure 2.3 below, the box entitled 'Risk assessment process' depicts the quantitative risk assessment stage. This stage involves an independent assessment of the hazards, exposure and risks associated with each specific pollutant being considered for a standard. While the risk assessment is a technical procedure based on scientific evidence, it may also take into account certain risk management policy issues, such as the precautionary principle and environmental equity. For example the precautionary principle is taken into account when deriving exposure-response function from overseas or toxicology data when no suitable Australian or epidemiology data are available. Environmental equity/justice issues are taken into account to ensure protection of vulnerable groups and sensitive individuals in the population.

The quantitative risk assessment results in a range of potential risk-based air quality guidelines for a specific pollutant. This is illustrated in Figure 2.3 below, in the box entitled 'Range of potential risk based guidance values'.

To derive an air quality standard with a statutory basis from the range of potential risk-based guidelines requires consideration of a number of other issues besides health, exposure and risk considerations. This is illustrated in figure 2.3, in the box entitled 'Policy considerations under the NEPC Act'.

Risk management involves a broader evaluation of the results of the risk assessment and takes into account the economic, social and environmental impacts of adopting standards. Impact statements and cost-benefit analysis are conducted. An impact statement may explore alternative approaches to risk management aside from setting air quality standards, such as exposure reduction strategies, that may result in better air quality outcomes. The evaluation must also take into account stakeholder views and consult with the public. Section 7 provides more information on the range of other considerations involved in setting air quality standards. In many cases the final standards are a balance of all of the issues and are set with an inherent level of human health risk associated with them (NHMRC 2006).

Figure 2.3 Framework for NEPC standard-setting process



2.5 RECOMMENDED FRAMEWORK FOR SETTING AIR QUALITY STANDARDS IN AUSTRALIA

The recommendations for the framework for setting air quality standards in Australia are:

- The process should be informed by a quantitative risk assessment approach, using the framework set out in Figure 2.3.
- The policy considerations set out in the NEPC Act, i.e. the precautionary principle and environmental equity considerations inform the scope of the technical risk assessment process.
- The risk assessment process must be separated from consideration of policy issues, and from the risk management processes that need to be considered in the final determination of a standard.
- All policy and decision-making steps must be transparent and explicit.

1. Refer to previously endorsed NEPC Methodology for prioritising pollutants

3 HAZARD ASSESSMENT—PRINCIPLES AND GUIDANCE

Individual research findings in themselves are not ready for direct input into a decision making or standards setting process. They usually come from a variety of scientific branches and need to be pre-processed before the overall information is useful for the standard setting process. Section 3 sets out the steps necessary to derive the specific information using the hazard assessment process. Section 3 commences with the weight of evidence approach providing guiding principles for assessing the quality and strength of the science used in the hazard assessment and for the final conclusion.

Hazard assessment incorporates two components of risk assessment — hazard identification and exposure response. It requires evaluation of literature from different fields to identify:

- key health effects and critical health outcomes
- exposure–response (or dose–response functions)
- sensitive groups in the population.

All international agencies that develop air quality guidelines and standards conduct hazard assessments that have the following features in common:

- an extensive literature search based on relevant key words and specific questions
- selection and evaluation of key epidemiological, controlled human exposure and toxicological studies to identify health effects associated with exposure to the pollutants of concern
- identification of susceptible or vulnerable groups within the population
- identification of exposure–response or no observed adverse effect level (NOAEL) or lowest observed adverse effect level (LOAEL)
- a weight-of-evidence analysis.

Both epidemiology and toxicology provide important information to guide hazard assessment. Epidemiological studies (Section 3.2.1) provide evidence of the impact of air pollution on populations at actual exposures experienced. In contrast, human toxicological (controlled human exposure) studies (Section 3.2.2) and animal toxicological studies (Section 3.2.3) are conducted on limited numbers of subjects (the most sensitive subgroups are generally not included) or on animals, and the exposures, although well defined, do not reflect real-world exposures. These types of study can provide evidence for plausible biological pathways, and therefore strengthen causal links between exposure and health effects. Evidence for an adverse effect from all relevant scientific areas needs to be combined through a weight-of-evidence process.

The following sections provide an overview of national and international approaches to identifying hazards and assessing exposure–response (or dose–response), and recommend an approach for standard setting in Australia.

3.1 WEIGHT OF EVIDENCE

As noted above, the evidence for adverse health effects from air pollution arise from a number of disciplines, including epidemiology, controlled human exposure studies and toxicology.

The weight-of-evidence process is widely used in decision making in public health. It overcomes the scientific uncertainty associated with any large body of evidence assembled in gaining knowledge of multi-factorial disease pathogenesis. It involves reaching a conclusion based on reasoning and probabilistic judgement, by using all available information and keeping in mind that absolute certainty about the causal relationship is elusive (Lucas 2005).

Weight-of-evidence approaches have been formulated by a number of regulatory and science agencies, sometimes for a specific class of agents such as carcinogens. For example, the International Agency for Research on Cancer (IARC) was the first organisation to develop a weight

of evidence scheme for cancer agents. A panel of international experts systematically evaluates the evidence of carcinogenicity, classifies each agent and publishes a summary of the evidence which includes the rationale used to support the agent's classification. IARC is an agency of the World Health Organization (WHO) (Muller 2002).

Although the IARC ranking continues to be highly respected, other agencies have developed similar ranking schemes. Of these, the one published by the US EPA (1986) is probably the most influential. The guidelines for cancer risk assessment finalised and released by the US EPA in 2005 (US EPA 2005a) recommend the use of a weight of evidence narrative that includes both a conclusion about the weight of evidence of carcinogenic potential and a summary of the data on which the conclusion rests. This narrative is a brief summary that *in toto* replaces the alphanumerical classification system used in EPA's 1986 cancer guidelines.

More general weight of evidence approaches have also been developed e.g. US EPA (US EPA 2005), the United States Center for Disease Control and Prevention (2004) and the WHO (WHO 2000b). In Australia, the NHMRC guidance document on health-based hazard assessment recommends a weight-of-evidence approach for setting ambient air quality standards (NHMRC 2006). The different organisations take a similar approach to weight of evidence, and the approaches provide a uniform structure for determining causality.

Where a range of evidence is available for a particular standard, it is important to be clear about how to determine the relative weights to be given to the different data sources. This is done by analysing the potential cause–effect relationships, the quality of the underlying studies, the consistency of results across studies and the biological plausibility of the exposure–response relationships. Each of these steps is discussed below.

The most direct evidence of a causal relationship between pollutant exposure and health effects comes from controlled human exposure studies. Such studies evaluate health effects linked to exposure under controlled laboratory conditions. Advantages of controlled exposure studies are that they enable an assessment of individual health response to a range of known single-pollutant exposures for fixed periods of time; hence, they reduce the potential for other unmeasured factors to affect an individual's health. Such studies are particularly expedient for studying the more proximal relationships between exposure and health, by examining the mechanisms of disease induction.

Disadvantages of controlled exposure studies are that they are limited by ethical considerations (e.g. the exclusion of the most sensitive individuals and the fact that only minor, short-term and reversible health effects can be investigated), are restricted to small groups that may not be representative of the general population and can generally only be used for short-duration exposures. Because of these issues, a lack of observable effects from controlled exposure studies does not necessarily mean that there is no causal relationship.

Epidemiological or observational studies of humans cannot explore the biological pathways of disease. To complement information derived from controlled exposure studies, population-based epidemiological studies provide estimates of the impact of exposures to these pollutants in the 'real world' (i.e. people breathe a mixture of pollutants, not just a single pollutant) and for the range of susceptible groups found in populations. These studies are also useful in deriving information about the impacts of longer exposures. However, they are limited by their ability to control for the range of other factors that may also affect health, and by the difficulty of accurately estimating exposures to the pollutant of interest.

Situations based on intervention such as closure or elimination of a pollution source allows epidemiological approaches to be used to compare health effects before and after the exposure

change. Examples of these types of studies, which have been considered to provide compelling evidence of causality, are mandated sulfur dioxide reductions in fuel in Hong Kong and the coal burning ban in Dublin, both of which resulted in subsequent improvements in health outcomes (Clancy et al 2002, Hedley et al 2002).

Overall, population-based randomised controlled studies are realistically not feasible, and the opportunities for intervention studies are rare. Therefore, other research designs, such as cross-sectional, time-series, case-control and cohort studies, all contribute valuable information to public health decision making, where a tradeoff between the internal precision of human clinical studies and the external precision of epidemiological studies has to be made.

Experimental animal data can complement clinical and observational data. In the absence of clinical and epidemiological human-based data, animal data may be sufficient to support a likely causal determination in cases where humans are assumed or known to respond similarly to the animals studied.

Much of the available health information related to air pollution comes from epidemiological studies that report statistical associations between exposure and health outcomes. Many of the health outcomes reported in these studies have complex etiologies, and depend on a variety of risk factors in addition to the environmental factors, such as age, genetic susceptibility, nutritional status, immune-competence, and social and economic factors.

Moving from association to causation involves eliminating alternative explanations for associations. By itself, an association cannot prove a causal relationship between exposure and disease. In relation to air quality standards, 'cause' explains a significant relationship between exposure to an air pollutant and an associated health effect, whereas 'association' is the statistical dependence among events, characteristics or other variables.

Inferring causation from epidemiologic studies requires consideration of uncertainties, particularly potential confounders. There are several ways to reduce the uncertainty in observed associations through statistical analyses; for example, through multivariate regression models and stratified analyses. Appropriate statistical adjustment for confounders requires identification and measurement of all reasonably expected confounders.

Confidence that results are not biased by measurement errors, including unmeasured confounders, is increased when multiple studies are conducted in various settings using different populations or exposures. Thus, multi-city studies that use a consistent method to analyse data from across locations with different levels of co-pollutants can provide insight on potential confounding in associations. Multivariate models are widely used to address confounding in epidemiologic studies, but such models need to be interpreted carefully when assessing effects of air pollutants.

Estimating the influence of exposure to an air pollutant on a health outcome requires two distinct levels of uncertainty to be considered (US EPA 2008):

- *model uncertainty* — that is, uncertainty about gaps in scientific theory required to make predictions on the basis of causal inferences
- *parameter uncertainty* — that is, uncertainty as to the statistical estimates within each model.

Assessment of model uncertainty involves consideration of the following policy-relevant issues:

- whether exposure to a particular pollutant causes the health outcome
- the set of confounders associated with exposure and health outcome
- which parametric forms best describe the relationships among exposure, confounders and outcome
- whether other forms of bias could be affecting the association.

Two steps are needed to address the policy-relevant issues noted above:

- *Step 1* — Determine the effects (if any) of the air pollutant under consideration on susceptible populations, given the total body of evidence.
- *Step* 2 — Determine at what levels of exposure health effects of concern occur.

The first step determines the weight of evidence in support of causation, and characterises the strength of any resulting causal classification. The second step includes further evaluation of the quantitative evidence about the exposure–response relationships and the levels, duration and pattern of exposures at which effects are observed.

Air quality standards provide a target for actions to improve air quality. These actions are implemented through jurisdictional air quality management policies and strategies. The basic principles of assessing weight of evidence for health effects and other aspects of hazard assessment should apply in setting standards. However, the key question is how to ensure that any standards that are developed are based on sound scientific evidence. The following section provides guidance on how to evaluate different types of studies to provide evidence for a link between air pollution and adverse health effects.

3.1.1 Bradford-Hill viewpoints

Although it is not possible to prove absolutely that any exposure causes (or does not cause) a given disease, a strong probability of causation provides a compelling basis for action. 'Proof' rests on an informed scientific interpretation of observations. This interpretation is guided by the 'viewpoints' set out in 1965 by the British medical statistician Sir Austin Bradford-Hill in his seminal paper on environmental cause of disease (Hill 1965). Bradford-Hill's approach has been re-appraised by Lucas & McMichael (2005) in the light of 40 years of epidemiological research. Bradford-Hill's viewpoints are:

- *strength of association* — which refers to the magnitude of the associations
- *consistency* — which refers to the extent of finding similar results in other studies
- *temporality* — which refers to the time sequence between exposure and health outcomes
- *exposure–response relationship* — which refers to evidence of increases in exposure causing increased risks in health effects
- *biological plausibility* — which refers to evidence of a plausible biological mechanism
- *coherence* — which refers to the assertion that it may weaken causal conclusion if the result of a study contradicts present substantive knowledge
- *strength of study design* — evidence from 'true experiments' is most compelling; for example, randomised control trials (RCTs) of exposure to gaseous pollutants have been performed with human volunteers; however, such experiments are often not feasible or ethical
- *analogy* — which refers to judging causality by observation of similar effects in relation to similar exposure
- *experiment* — which refers to evidence-based on randomised experiments.

In addition to these viewpoints, bias and confounding have to be considered as possible alternative explanations of the relationship between exposure and effect. Bias may occur due to systematic errors in measuring exposure or health, or may be introduced during selection of the study population. Confounding may occur when risk factors (known or unknown) are closely related to both the exposure under investigation and the health effects. Confounding by co-pollutants is an ongoing issue in air pollution epidemiology, and has to be considered when evaluating the evidence.

To make it easier to interpret air pollution and health data, the US EPA has modified the Bradford Hill viewpoints for use with a broader array of data, including epidemiological studies, human clinical studies and animal toxicological studies, as well as in vitro data. Table 3.1 shows these

modifications, which are recommended as the basis for assessing causality when setting air quality standards in Australia.

Table 3.1 Aspects to aid judging causality (adapted from the US EPA 2008)

1. ***Consistency of the observed association***	An inference of causality is strengthened when a pattern of elevated risks is observed across several independent studies. The reproducibility of findings constitutes one of the strongest arguments for causality. If there are discordant results among investigations, possible reasons — such as differences in exposure, confounding factors and the power of the study — are considered.
2. ***Strength of the observed association***	The finding of large, precise risks increases confidence that the association is unlikely to be due to chance, bias or other factors. A modest risk, however, does not preclude a causal association and may reflect a lower level of exposure, a pollutant of lower potency, or a common disease with a high background level.
3. ***Specificity of the observed association***	There is increased inference of causality if one cause is associated with a single effect or disease. The US EPA now considers this to be one of the weaker guidelines for causality; for example many agents cause respiratory disease, and respiratory disease has multiple causes. The ability to demonstrate specificity under certain conditions remains, however, a powerful attribute of experimental studies. Thus, although the presence of specificity may support causality, its absence does not exclude it.
4. ***Temporal relationship of the observed association***	A causal interpretation is strengthened when exposure is known to precede development of the disease.
5. ***Biological gradient (exposure–response relationship)***	A clear exposure–response relationship (e.g. increasing effects associated with greater exposure) strongly suggests cause and effect, especially when such relationships are also observed for duration of exposure (e.g. increasing effects observed following longer exposure times). There are many possible reasons why a study may fail to detect an exposure–response relationship. Thus, although the presence of a biological gradient may support causality, the absence of an exposure–response relationship does not exclude a causal relationship.
6. ***Biological plausibility***	An inference of causality tends to be strengthened by consistency with data from experimental studies or other sources demonstrating plausible biological mechanisms. A lack of biological understanding, however, is not a reason to reject causality.
7. ***Coherence***	An inference of causality may be strengthened by other lines of evidence that support a cause-and-effect interpretation of the association. For instance, similar findings between clinical and animal studies, or closely related health effects, which are expected to be associated with exposure, are observed together. The absence of other lines of evidence, however, is not a reason to reject causality.
8. ***Experimental evidence (from human populations)***	Experimental evidence is generally available from human populations for the criteria pollutants. The strongest evidence for causality can be provided when a change in exposure brings about a change in adverse health effect or disease frequency in either clinical or observational studies (e.g. intervention studies).
9. ***Analogy***	Structure activity relationships and information on an agent's structural analogues can provide insight into whether an association is causal. Similarly, information on mode of action for a chemical, as one of many structural analogues, can inform decisions regarding likely causality.

The aspects given in Table 3.1 provide a framework for assessing the evidence, but they do not lend themselves to consideration in terms of simple formulas or fixed rules for drawing conclusions on causality (US EPA 2008). The goal of taking these considerations into account is to produce an objective appraisal of the evidence. In particular, the absence of one or more of the aspects does not automatically exclude a study or a body of evidence from consideration.

The evaluation of epidemiological literature and assessment of causality should be completed by experts familiar with the subject matter and the required methods (Rothman and Greenland 2005). The WHO has based its conclusions on the critical scientific judgment of a wide range of scientists working in various disciplines related to the assessment of impacts of air pollution on health. Experts are used in preparing technical background information (hazard identification) and for drawing conclusions on the assembled body of evidence. Similar processes are used in the United Kingdom and the United States. An independent peer review process is generally used to assess the findings of the expert groups.

Given that the weight-of-evidence assessment is an incremental logical reasoning process, whereby critical scientific appraisal is used to draw conclusions on the overall evidence, the process should be well documented and transparent.

3.1.2 Evaluating the weight of evidence — coming to a conclusion

A weight-of-evidence assessment takes into account all lines of evidence including human studies (clinical and epidemiological), animal studies and in-vitro studies. The separate judgments are then integrated into a qualitative statement about the overall weight of the evidence and causality.

Further issues to be considered in evaluating the weight of evidence for an effect relate to characterising exposure and risk to populations; that is, determining at what levels health effects occur. Questions that need to be considered are:

- What is the exposure–response relationship?
- Under what exposure conditions (dose or exposure, duration and pattern) are effects seen?
- What population groups appear to be affected or more susceptible to effects?

For the purpose of hazard assessment for setting air quality standards in Australia, the approach developed by the US EPA (2008) has been adapted to guide the weight of evidence evaluation for determination of causality (Table 3.2).

Table 3.2 Weight of evidence for causal determination (adapted from US EPA 2008)

Sufficient to infer a causal relationship	Evidence is sufficient to conclude that there is a causal relationship between relevant pollutant exposure and the outcome. Causality is supported when an association has been observed between the pollutant and the outcome in studies in which chance, bias and confounding could be ruled out with reasonable confidence. Human clinical studies provide the strongest direct evidence for causality. Causality is also supported by findings from epidemiologic 'natural experiments' or observational studies supported by other lines of evidence. Generally, determination is based on multiple studies from more than one research group.
Sufficient to infer a likely causal relationship (i.e. more likely than not)	Evidence is sufficient to conclude that there is a likely causal association between relevant pollutant exposures and the outcome. That is, an association has been observed between the pollutant and the outcome in studies in which chance, bias and confounding are minimised, but uncertainties remain. For example, observational studies show associations, but confounding and other issues are difficult to address and/or other lines of evidence (human clinical, animal or mechanism of action information) are limited or inconsistent. Generally, determination is based on multiple studies from more than one research group.
Suggestive, but not sufficient to infer a causal relationship	Evidence is suggestive of an association between relevant pollutant exposures and the outcome, but is weakened because chance, bias and confounding cannot be ruled out. For example, at least one high-quality study shows an association, while the results of other studies are inconsistent.
Inadequate to infer the presence or absence of a causal relationship	The available studies are inadequate to infer the presence or absence of a causal relationship. That is, studies are of insufficient quality, consistency or statistical power to permit a conclusion regarding the presence or absence of an association between relevant pollutant exposure and the outcome. For example, studies that fail to control for confounding or that have inadequate exposure assessment, fall into this category.
Suggestive of no causal relationship	The available studies are suggestive of no causal relationship. That is, several adequate studies, examining relationships between relevant population exposures and outcomes, and considering sensitive subpopulations, are mutually consistent in not showing an association between exposure and the outcome at any level of exposures. In addition, the possibility of a small elevation in risk at the levels of exposure studied can never be excluded.

3.2 HAZARD IDENTIFICATION

Hazard identification involves determining what types of adverse health effects may be caused by exposure to the pollutant, and how quickly the adverse effects might be experienced. This involves reviewing available epidemiological and toxicological data, taking into account the strengths and weaknesses of the information being reviewed.

The specific objectives in the hazard identification process are to identify all available studies about the effects of the pollutant in question using a systematic search of all relevant databases and predetermined inclusion criteria and timeframe; the search should be based on specific questions posed by the review. The evidence is then evaluated using guiding questions to explore the scientific validity of the individual studies (see Section 3.1).

A comprehensive bibliographic search would include the following:

- definition of the key policy questions relevant for the review
- an explicit search strategy, including identification of key words
- involvement of qualified searchers (e.g. librarians and trained investigators)
- searching of bibliographic databases and an effort to include all available studies

Other methods that could be included, depending on the search protocol, include hand searching of journals and inclusion of abstracts and unpublished data (including writing to authors of published data). If established, an expert group (as discussed in Section 3.5) would need to decide the search strategy and criteria for quality. All potentially relevant studies should be identified at the first stage of the expert group's work. However, depending on the types of exposure evaluated, whole categories of studies may be excluded in the second step, on the basis of quality criteria.

The following are the key policy-relevant questions for the review of the scientific evidence:

- Has new information altered the scientific support for the occurrence of health effects following short and/or long-term exposure to levels of air pollutants found in the ambient air in Australian cities?
- What do recent studies focused on source environments tell us about health effects of air pollutants?
- At what levels of exposure to air pollutants do health effects occur?
- Has new information altered conclusions about the plausibility of adverse health effects caused by exposure to air pollutants?
- To what extent have important uncertainties been identified and addressed?
- What are the relationships between short and long-term exposures to air pollutants and adverse health effects?

The first step in reviewing and updating the current knowledge base of health effects of air pollution for the purpose of conducting a hazard assessment should be the appraisal of current reviews conducted by national and overseas agencies. These reviews are often comprehensive and are based on scientifically agreed approaches (see Section 3.1), making it easier to evaluate the evidence. What are important in these reviews, especially when they originate from standard-setting agencies, are the key studies and exposure response estimates that were thought to be important in coming to a conclusion or setting a standard. A comparative overview of the most salient points of the final weight of evidence findings and the key studies should be prepared.

The second step should involve a search of studies that may have become available since the publication of recent reviews, including a search for all relevant Australian studies. The conclusion derived from recent reviews should then be compared and discussed in the light of the new emerging evidence.

In evaluating the evidence for adverse health effects arising from exposures to air pollutants, it is important to consider the types of studies. In investigating the health effects of air pollution, it is important to integrate the following types of study. Each approach has specific strengths and weaknesses in evaluating the human health effects of air pollution.

- *Epidemiological studies* examine the relationship between air pollution exposure and health effects in the community. They can investigate acute or chronic (long-term) effects. Accurate estimation of exposure to a pollutant is usually difficult. Exposures are generally estimated from fixed monitoring sites, and many pollutants occur as components of complex mixtures. The extent of potential confounding factors (e.g. cigarette smoking and health status), time considerations in air pollution effects (e.g. lags and latencies), individual variation in air pollution exposure, and exposure misclassification cause uncertainty in any observed associations with the health outcome.
- *Controlled human exposure studies* (or chamber studies) investigate mechanisms of injury, and permit strict control of the exposures and the characteristics of the exposed persons. Ethical and practical considerations limit the use of controlled human exposure studies, and chronic effects cannot be readily addressed. Despite these constraints, such studies have contributed significantly to quantification of the relationship between respiratory function and air pollution (e.g. the link between lung function and ozone exposure).
- *Toxicological studies* of animals can evaluate mechanisms of injury by pollutants using methods that cannot be applied to human subjects. Uncertainty is introduced when animal models are extrapolated to humans.

3.2.1 Epidemiological studies

A large and relevant contribution to the hazard assessment for air pollutants comes from the field of epidemiology, which explores relationships between exposures and health outcomes in a human population setting. Results from epidemiological studies play a central role in the policy area and particularly in standard setting, because the data are based on the response of human populations (including the most sensitive groups) to real exposure scenarios.

The most commonly used epidemiological study designs for investigating the effects of air pollution on health for the purpose of standard setting, discussed below, are:

- time-series
- case-crossover
- cross-sectional
- case–control
- cohort
- panel
- intervention.

Originally, a strict distinction was made between ecological and observational studies, with the unit of observation (exposures and health outcomes) being based on either aggregated group-based data (ecological studies) or on individual (observational studies) data. In air pollution epidemiology, this distinction has been blurred, and aggregated exposure data are used across all epidemiological study designs (Samet and Jaakkola 1999). In air pollution studies, the exposure of participants is generally not individually observed, and the concentrations of air pollutants at one or more stationary air monitors are used as a proxy for individual exposure.

3.2.1.1 Time-series studies

Aggregated information at the population level is used in time-series studies to describe both exposures and health effects. Health outcomes are obtained from databases that routinely collect daily data on mortality and morbidity, (including hospital and emergency department admissions), using coding by diseases, based on the International Classification of Diseases (ICD) codes.

Worldwide, the time-series study design has been relied on for assessment of short-term changes of health outcomes such as morbidity and mortality daily events in relation to air pollutant

measurements. Health outcomes data are gathered on a daily basis and are analysed in relation to daily air quality levels relevant to the community location. Time-series studies are particularly informative because they allow the study of associations between *changes* in outcomes and *changes* in exposure indicators preceding or simultaneous with the outcome. Due to this temporal relationship, the results from time-series studies are supportive of a causal interpretation, even when both the outcome (e.g. the number of non-accidental deaths in a city during a day) and the exposure (e.g. daily air pollutant concentration) are community indices. Furthermore, due to its temporal associations, this study design is also less vulnerable to potential confounding, especially those confounding risk factors that change over time such as weather, other air pollutants and episodes of major illnesses in the community (e.g. influenza), because they can be controlled for during statistical analysis (RATF 2000).

For time-series studies, it has long been realised that there is an association between temperature (and, more generally, weather) and various acute health outcomes including daily mortality, daily hospital admissions and daily hospital emergency department attendance. For example, extremes of temperature, both hot and cold, can cause increased deaths. Because of the potential for weather to confound the air pollution-health outcome relationship, its effects must be controlled in the analysis of time-series studies so that the independent effects of air pollution on the chosen health outcome can be determined. Other variables that can confound results are the presence of infectious diseases, aeroallergens and other chemicals. Other types of confounders such as socioeconomic or individual life style factors (e.g. smoking and diet) cannot be controlled for, because this individual information is not available at an aggregate level. As these confounders do not change significantly over the timeframe of these studies, confounding by these types of variables is not a major issue in time-series analysis.

3.2.1.2 Case-crossover studies

The case-crossover design is suited to the study of a transient effect of an intermittent exposure on the subsequent risk of a rare acute-onset disease a short time after the exposure has occurred. The case-crossover design has been used to study the association between short-term air pollution exposure and acute health effects using community and individually based datasets of exposure and health effects.

The principle of the analysis is that the exposures and rates of health effects just before the event are compared with the distribution of exposures estimated from some separate time period. This distribution of exposures is assumed to be representative of the distribution of exposures for those individuals while they are at risk of developing the outcome of interest. With this approach, cases serve as their own controls, and confounders that do not change over time are controlled by design. Matching techniques can be used to test for confounding, by variables that change over time such as season, temperature or co-pollutants (Janes et al 2005). The potential of the case-crossover design to deal with confounding in this way has made it a popular alternative design to the time-series approach.

A disadvantage of the case-crossover design, however, is the potential for bias due to time trends in the exposure time series. Since case-crossover comparisons are made between different points in time, the case-crossover analysis implicitly depends on an assumption that the exposure distribution is stable over time (stationary). If the exposure time-series is non-stationary and case exposures are compared with referent exposures systematically selected from a different period in time, a bias may be introduced into estimates of the measure of association for the exposure and disease. Various strategies can be used to reduce bias and confounding during design and analysis of such studies (Navidi 1998, Janes et al 2005).

3.2.1.3 Cross-sectional studies

Cross-sectional studies are used to compare the prevalence of health effects of air pollution across several locations at a given point in time. The objective is to assess differences in the prevalence of health effects in relation to differentials in exposure at different locations. Cross-sectional studies that are based on a random sample of the members of a given population are often described as surveys (US EPA 1996b, Vedal 1997). These studies have an advantage over time-series studies in that they capture individually based information on acute and chronic effects in relation to air pollution measurements, and they can gather data on a large number of potentially confounding factors. Because exposure and health outcomes are gathered as a snapshot in time, cross-sectional studies cannot be used to assess causality; instead, they are used to assess hypotheses about associations between diseases and air pollution. The Australian Children's Health and Air Pollution Study (ACHAPS), conducted by the University of Queensland and Woolcock Institute, is an example of a cross-sectional study. This study was commissioned by EPHC to inform the review of the standards in the AAQ NEPM. The study investigates sensitive respiratory outcomes such as lung function and respiratory symptoms from a cross-section of children recruited across all major cities in Australia, in relation to local air pollution levels.

3.2.1.4 Case–control studies

Case–control studies are retrospective, in that exposure is determined after the health endpoint occurs (common in occupational health studies). Individuals with a given disease (cases) are compared with controls without the disease in relation to their possible exposures in the past. This study design is prominently used in assessing the effect of possible exposures on health outcomes such as cancer in the occupational environment.

Population-based case–control studies also played a significant role in identifying relationships between birth defects and air pollution (Gilboa et al 2005, Ritz et al 2007). In these studies, cases with identified birth defects were retrospectively compared to children born without birth defects on the basis of their foetal exposure to air pollution.

3.2.1.5 Cohort studies

In cohort studies, subjects are selected based on exposure status and are followed to monitor the development of a specific health endpoint (US EPA 1996b, Vedal 1997). Cohort studies can be conducted prospectively or retrospectively. In a prospective cohort study, exposure status is determined from current or historical records, and the subjects are followed to monitor the development of disease. For prospective cohort studies, extensive exposure assessment can be undertaken.

If the individuals selected to participate in a prospective cohort study are representative of air pollution exposures across different communities, the effects of individual risk factors can be separated from exposure to air pollution. This epidemiological design also allows the evaluation of cumulative exposure to air pollution over several years, whereas acute effects study design only allows assessment of effects of short-term exposure changes (US EPA 1996b).

These strengths of prospective cohort studies are generally reduced if only occasional air pollution measurements are available, so that only crude exposure comparisons across cities or regions can be made. The disadvantages of this design are the potential difficulty and high cost of implementation. The follow-up of study populations over extended periods of time is difficult. Large numbers of subjects are required if rare diseases are to be considered. This study design generally has weak power to measure interactions.

The best way to measure long-term health impacts of air pollution is by using a cohort design that follows a large cross-section of people over time. These studies allocate aggregated air pollution levels of the nearest city or regional level to the recruited individuals, and over time they observe

the development of individual ill health and deaths at various follow-up times. Individual data on other factors, such as nutritional status, smoking and family history of disease are also gathered. The cost of individual follow-up in such studies is high; therefore, only few cohort studies are available. The study design has a clear time relationship between exposure and health outcome, and individual data can be used to assess the effect of important confounders. These features make this study design relevant for assessment of causality. Two commonly cited studies, the Harvard Six Cities study and the American Cancer Society study, both undertaken in the United States, have measured the effects of exposure to particulate matter on mortality.

Prospective cohort studies have also been used successfully to evaluate the acute effects of time-varying exposures to single air pollutants on daily reports of symptoms and changes in pulmonary function.

3.2.1.6 Panel studies

Panel studies follow up day-to-day health changes of a group of participants in relation to daily changes in air quality measures. Examples of panel studies include those in sensitive subgroups such as children, asthmatic children and elderly people; they assess daily symptoms, lung function changes and use of medication using a daily diary. Panel studies are essentially a cohort type, albeit with a short-term follow-up. A prominent example is the Pollution Effects on Asthmatic Children in Europe (PEACE) multi-panel study, which conducted panel studies in children from 14 different centres in Europe, collecting respiratory outcomes in relation to daily changes in air pollution (Roemer et al 1999).

Panel studies have been widely used in investigating the health effects of the criteria pollutants on sensitive groups, including people in high-exposure groups, such as children on summer camps (Kinney et al 1996), hikers (Korrick et al 1998) and outdoor workers (Brauer et al 1996). These studies have focused on individuals with asthma, and usually involve small groups of subjects. The value of these studies is that they provide information on exposure–response at an individual level, and allow for control of individual level factors that may influence the health outcome of interest. Panel studies have been of particular importance in identifying significant associations between ozone and respiratory symptoms, mainly in asthmatic children (Romieu et al 1996, Gielen et al 1997, Hiltermann et al 1998, Gold et al 1999, Desqueyroux et al 2002a and b, Ross et al 2002).

3.2.1.7 Intervention studies

Intervention studies test the hypothesis of a cause-effect relationship by modification of a supposed cause and by measuring the effect of the change. For the purpose of studying health effects of air pollution at the population level this experimental approach is usually not feasible unless an independent opportunity of change arises. Interventions such as the ban on the sale of coal in Dublin and the reduction of sulfur in fuel in Hong Kong provided a valuable opportunity to demonstrate that improved air quality is associated with significant reductions in mortality (Clancy et al 2002, Hedley et al 2002). These types of environmental intervention studies add extra weight to a causal explanation of the health effects of air pollution.

3.2.2 Controlled human exposure studies

Population-based epidemiological studies cannot control every aspect of exposure, health outcomes and the study population. In contrast, experimental studies in humans enable control of these factors. The essential elements include a control group and one or more exposed groups.

Chamber studies can play a vital role in assessing effects of single or combined air pollutants on objectively measured health outcomes; for example, in pulmonary function and bronchial hyper-responsiveness testing. These studies, together with epidemiological evidence, have helped to provide useful information for setting standards and guidelines for ozone.

Toxicological experiments on human subjects are generally conducted in controlled exposure chambers, where endpoints include measurements of respiratory and heart function, and analysis of bronchoalveolar lavage products. In controlled human exposure studies the most sensitive populations are not studied for ethical reasons, and these studies are generally limited to the examination of health effects that are short term, mild and reversible. The subjects are often healthy young male adults, who may have a different response to more sensitive individuals (e.g. older adults and children). For these reasons, clinical studies of the effects of pollutants — for example, particulate matter in potentially susceptible individuals, such as those with mild asthma, chronic obstructive pulmonary disease or heart and blood vessel disease — are highly desirable. In other cases, such as the adverse health effects of polycyclic aromatic hydrocarbons, toxicological research has an important role because few epidemiological data are currently available.

Randomised Control Trials are not limited to the laboratory setting. For environmental purposes, they have been used in community settings for determining health effects from indoor sources of exposure; for example, from combustion of bio-fuels and unflued gas heating (Pilotto 2004, Diaz et al 2007). RCTs provide a high level of causal credibility, but are unrealistic for studying health effects in relation to ambient air pollution in the general population. The advantage of RCTs is the high internal validity of the results which provides a high degree of confidence in the findings in relation to the study population. The disadvantage is that the results may not be valid for the general population and for subgroups of the population, who may not be represented in the clinical trials population. For example, children and adults with particularly severe chronic diseases (e.g. asthma) are rarely included because of ethical issues. Small sample sizes and tightly controlled exposures may not be representative of realistic scenarios. Also, interaction with other pollutants, weather and other seasonal confounders may play a role, and cannot be taken into account in a strict experimental setting.

3.2.3 Animal toxicological studies

Advantages of toxicological studies are that they:

- can be used to study the biological mechanisms underlying the symptoms and health effects observed in epidemiological studies
- often focus on the early events in the pathway to clinically evidenced disease
- can further investigate detailed exposure–response relationships when observational studies are not appropriate
- can reduce uncertainties over the causative role of a single pollutant which normally occurs in a complex ambient mixture
- can be used to study pollutant mixtures, which may display complex dose–response relationships and exhibit latency, as well as cumulative, non-additive and even synergistic effects.
- can be used for mechanistic studies in molecular, cellular and animal models, where gross effects on organs such as the lungs and heart, physiological responses such as inflammation, or direct effects such as genotoxicity can be described and interpreted; such studies include the absorption of particles from the lungs into the blood stream and subsequent distribution to sensitive organs or tissues, and the impact of particles on blood clotting pathways.

Inhalational toxicological studies using whole animals (usually rodents) may be particularly informative for setting air quality standards. A comprehensive monograph edited by Salem and Katz (2006) discusses the methodology, interpretation and extrapolation of inhalational toxicology studies in animals.

Among the drawbacks of using whole animal studies as indicators of potential human health effects of air pollutants is interspecies extrapolation. As the respiratory anatomy and physiology of rodents is not a particularly close match to humans, the achieved dose at sensitive sites can vary

widely between rodents and humans. Additionally, humans display much more inter-individual variability in some physiological responses than litter-mates of an inbred rodent strain. The development of animal models for human disease or susceptibilities such as asthma, high blood pressure, and chronic lung inflammation helps to overcome some of these shortcomings, and animal studies continue to be the main method used for predicting adverse health outcomes in humans (Dybing and Totlandsdal 2004).

3.2.4 Weighting of evidence for hazard identification

Toxicology and epidemiology can each add to the weight of evidence for the health effects of air pollutants. The absence of study results from either field does not indicate a weak and uncertain relationship, as has been demonstrated above. The evidence for the health effects of air pollutants comes from a number of different types of studies; however, the question of how to weight the evidence to ensure that the selection of health outcome and dose–response data is valid for the basis of setting air quality standards is a key issue.

An important aspect of the assessment for causality is the existence of biological evidence for a plausible relationship between health effects and exposure. This information comes largely from toxicological data and controlled human exposure studies, where such studies demonstrate underlying physiological processes that can explain health outcomes measured during epidemiological investigations.

For particles, PM_{10} and $PM_{2.5}$, the evidence for health effects has come largely from epidemiology. At the time of setting the first standards for these pollutants, little was known of the underlying biological pathways for PM_{10} and $PM_{2.5}$ although epidemiological studies showed a population effect (Lippmann et al 2000). Since that time, controlled human exposure and animal toxicological studies have provided evidence for biological mechanisms, supporting the results of the epidemiological studies. The findings from both toxicological and epidemiological studies using concentrated ambient particles have supported the evidence of biological plausibility of particles being causally involved in inflammatory respiratory and systemic processes. For ozone, there was a significant body of information from controlled human exposure and animal toxicological studies before epidemiological studies showed effects at the population level.

In contrast to the criteria pollutants, the potential health effects of air toxics range from respiratory irritation, central nervous system effects, reproductive and developmental effects through to cancer. Most of the evidence on the health effects of these pollutants arises mainly from occupational studies and animal toxicological studies. Historically, air toxics have not been routinely monitored in Australia, and no extensive databases are available that would allow researchers to conduct population-based studies that link ambient levels of air toxics to hospital admissions or mortality data (as has been done for the criteria pollutants). For carcinogenic substances, the latency period between exposure and onset of cancer can be long, which again makes population-based studies difficult. Studies undertaken in occupational settings are widely used to assess the risk posed by exposure to air toxics. However, these studies have been undertaken on workers who have been exposed at much higher levels of a pollutant than would be experienced by the general population. In extrapolating the results of these studies to assessing the risk to the general population, uncertainty factors must be applied to account for the differences in exposure levels. Further uncertainty factors must be applied to account for sensitive groups within the population, such as older adults, people with existing diseases and children, who would not be part of the occupational study population. Uncertainty factors also need to be applied to account for differences in exposure duration; for example, 8 hours a day, 5 days a week over a 40-year working lifetime for a worker compared to potentially 24 hours a day, 7 days a week over a 70-year lifetime for the general population. Guidance on the use of uncertainty factors is found in Sections 5.5 and 5.6.

Another significant source of data on the health effects of air toxics comes from animal toxicological studies. As with the occupational studies, these studies are conducted at concentrations of a pollutant much higher than would be experienced by the general population from environmental exposures. One key attribute of these studies is that the exposures are controlled and the issue of co-exposure to other pollutants, which can be a confounder in occupational studies, can be avoided. These studies provide important information and can be used as the basis for air quality standards, with uncertainty factors applied appropriately. The uncertainty factors are applied to account for animal-to-human extrapolation; differences in exposures; and sensitive groups within the population (see Sections 5.5 and 5.6 for further guidance).

The NHMRC (2006), states that it is important to integrate toxicology and human studies in assessing pollutants. Data from all scientific areas are complementary and should not be evaluated in isolation from each other. The organisation recommends that the usefulness of studies for standards setting be assessed on an individual basis and notes that it will depend on the nature of the exposure, the adverse health effect and the availability of data.

3.2.4.1 Use of epidemiological studies for setting air quality standards

Since the early 1990s, the results of epidemiological studies have been used extensively for setting air quality standards, due to the strength of epidemiology in studying long-term (cohort) and short-term health outcomes (time-series, panel studies). Positives attributes of epidemiology include the ability to:

- study large populations and extrapolate the results to the general population
- study populations under realistic exposure scenarios relevant to the population in question
- obtain risk estimates for quantification of mortality and morbidity in identified populations
- study a wide range of significant diseases in the general population and in sensitive subpopulations
- study populations without the ethical restraint placed on human experimental studies
- investigate the effect of exposures on health; for example, on quality of life and on socioeconomically deprived populations with pre-existing poor health status.

For the criteria pollutants, a large amount of the evidence for health effects, especially the evidence for an association between these pollutants and mortality or hospital admissions, comes from time-series studies.

As mentioned, epidemiological studies focusing on environmental factors have been used extensively in setting air quality standards, and their principal advantages derive from their direct relevance to human health and actual population exposures. They do, however, have limitations, including:

- exposure and outcomes being misclassified, which generally leads to an underestimation in risk estimates
- in large populations, substantial effects in small numbers of sensitive individuals possibly being swamped by a lack of effect in the majority
- confounding by closely linked co-exposures.

The findings of epidemiological studies have to be appraised carefully, because the size of the statistical associations between environmental exposures and health outcomes is usually small. This association may include other factors that may simultaneously affect the health outcomes under consideration. For example, possible simultaneous health effects of co-pollutants illustrate the confounding of the relationship between pollutants and health outcomes.

The multi-factorial nature of disease aetiology may make it hard to separate the contributions of individual pollutants. An appraisal of all of the evidence, including toxicological information about underlying mechanisms, is necessary for informed decision making.

Data from epidemiological studies, controlled human studies, and animal studies are complementary, and they should be considered together in identifying hazards, determining which studies are 'important' and analysing quantitative data.

3.2.4.2 Use of controlled exposure studies for setting air quality standards

Controlled human exposure studies are also used widely in setting standards. They have some advantages over population-based epidemiological studies, in that confounding is eliminated due to well-defined exposures and health outcomes. Any unknown confounders are distributed equally in the exposed and unexposed group. Also, misclassification of exposure and health outcomes is avoided by use of precisely known exposure concentrations. Accurate exposure–response relationships can be obtained for the study group, and this type of study can be used to test for effect thresholds in study populations. In addition, more sensitive health outcomes can be observed and measured (pulmonary function testing).

The main limitations of controlled human exposure studies for setting standards are that study subjects are restricted to healthy individuals or those with mild or episodic conditions, and real-life exposures cannot be replicated. This means that it may not be possible to generalise the study results to the broader population. In addition, the studies are restricted to investigating mild and reversible health outcomes, and cannot generally be used to evaluate the potential of a chemical to cause chronic disease. Chance variations are also possible due to limitations in sample size.

For the purposes of setting air quality standards, the selection of studies of controlled human exposures to air pollution for evaluation should emphasise those studies that:

- investigate potentially susceptible populations (e.g. mild asthmatics), and particularly studies that compare responses in susceptible individuals with those in healthy controls
- determine exposure-response functions
- investigate exposure to the pollutant of concern separately and in combination with other pollutants
- include control exposures to filtered air
- have sufficient statistical power to assess findings.

3.2.4.3 Use of animal toxicological studies for setting air quality standards

Animal toxicology studies have been used to investigate all of the criteria air pollutants and many of the air toxics as well. Such studies provide an important database of evidence for the health effects of air pollution, and can provide information on biological mechanisms and plausibility for the effects observed in epidemiological studies. The main advantage of animal studies is the ability to control exposures. Studies are usually designed such that the highest exposure results in measurable adverse effects. While ethical constraints limit the extent to which humans can be deliberately exposed to pollutants, it is possible to design ethically acceptable animal studies that allow valid adverse-effect data to be collected (NHMRC 2006). In addition, adverse health effects can be monitored through a combination of observation and measurements while the animal is alive (e.g. through blood chemistry, haematology, urinalysis and lung function) and through necropsy examination of tissues. The latter measurements are particularly important to identify target organs for pathological damage (e.g. liver, brain) and are of critical importance if cancer is a suspected endpoint.

The most obvious limitation in using animal studies is the uncertainty in extrapolating the findings to humans. This is generally addressed by the use of uncertainty factors (or 'safety factors'). The

default values for uncertainty factors take into account potential variations in sensitivity within and between species. The use of such uncertainty factors generally adds a degree of conservatism to the extrapolated data. Uncertainty may relate to:

- the route via which the chemical enters the animal's system
  - studies are often designed to expose animals (usually rodents) to controlled concentrations in exposure chambers for significant periods of time (including lifetime exposures)
  - where whole-body exposures occur, there may be some uncertainty about the relative extent to which systemic absorption occurs via inhalation versus ingestion, particularly for particles, where feed contamination and ingestion via grooming may confound the analysis
  - studies may be designed with nose-only exposure, using suitable air delivery systems, but the degree to which the animals must be restrained generally limits the duration and frequency of the exposures
- differences in respiration rate, tidal volumes, anatomy and oro–nasal breathing patterns
- the extent to which deposition on the surfaces of the delivery and exposure equipment may reduce the actual exposure concentrations.

Unless the study includes careful observation, it may be difficult to assess sensory deficits and adverse effects on behaviour. Headache, nausea and mucous membrane irritation (eye and respiratory passages) are likely to be of particular significance in humans exposed to air pollution, but they are difficult, if not impossible, to assess in animal studies. Animal studies usually employ high levels of exposure, where even the lowest level exceeds concentrations likely to be encountered in ambient air. Various mathematical approaches have been used to extrapolate dose response downward from the experimental range. Care must be taken to assess whether effects observed at high doses in animals are relevant to lower exposures. There are many instances in the literature where specific mechanisms of toxicity only become operational at high doses. A breakpoint in the dose–response curve is usually a clue to the possibility that the response seen in animals may be concentration dependent.

Criteria for selecting animal toxicological studies for inclusion in a review include a focus on studies conducted at levels within about an order of a magnitude of ambient concentrations for the pollutant of concern, and studies that approximate expected human exposure conditions in terms of concentration and duration. Studies that elucidate basic biological responses and mechanisms of action or susceptibility — particularly if conducted under atmospherically relevant conditions — should be emphasised whenever possible. These studies are useful in evaluating the 'plausibility' of epidemiological and other studies.

3.2.5 Overseas approaches

Internationally, the results of epidemiological studies are used widely as the basis of air quality standards and guidelines. Time-series studies are relied on for exposure–response data used in setting standards. Studies considered are those that relate adverse health outcomes in humans to measures of ambient air quality. Studies of occupational exposures (typically much higher than ambient levels) are not usually considered in the development of standards for the criteria air pollutants.

In the United States, the NAAQS are based almost solely on epidemiological data. For the hazard assessment, studies are selected that examine both long-term and short-term effects. For particles, studies that have examined different size fractions, or have analysed health effects of specific components of the particles or indicators of particle sources, are considered. Studies that have evaluated a range of health endpoints and sensitive populations are used, including those that assess health endpoints and populations not previously considered (e.g. diabetics, low socioeconomic groups).

Studies are also evaluated on whether multiple pollutants have been analysed, or other approaches have been used to address issues relating to potential confounding of effects or effect modification. Studies that have addressed important methodological issues (e.g. lag structure, model specification, thresholds and mortality displacement) are also identified.

A range of questions are considered in assessing the relative scientific quality of epidemiological studies and to assist in the interpretation of their findings:

- Was the quality of the air monitoring data used sufficient to allow for meaningful characterisation of geographic or temporal differences in study population pollutant exposures in the range of pollutant concentrations evaluated?
- Were the study populations well defined and adequately selected so as to allow for meaningful comparisons between study groups or meaningful temporal analyses of health effect results?
- Were the health endpoint measurements meaningful and reliable, including consistency in obtaining dependent variable measurements and clear definition of diagnostic criteria used?
- Were the statistical analyses used appropriate, and properly performed and interpreted, including accurate data handling and transfer during analysis?
- Was likely important confounding or covarying factors adequately controlled for or taken into account in the study design and statistical analyses?
- Were the reported findings internally consistent, biologically plausible and coherent in terms of consistency with other known facts?

These guidelines provide benchmarks for judging the relative quality of various studies and for focusing on the highest quality studies in assessing the body of epidemiologic evidence. Of most importance in selecting studies are those that provide useful qualitative or quantitative information on exposure–response relationships for health effects in humans associated with ambient levels of pollutants currently likely to be encountered in the United States.

Epidemiological studies are selected that address a wide range of health endpoints. These include;

- mortality (total non-accidental and cause specific) — considered to provide the most unambiguous evidence related to a clearly adverse endpoint
- cardiovascular and respiratory hospital admissions and emergency room attendances
- medical visits
- reports of respiratory symptoms
- self-medication in asthmatics
- changes in lung function
- changes in cardiovascular physiology or functions (e.g. heart-rate variability) and blood coagulation
- low birth weights, pre-term deliveries, etc.

All susceptible groups identified in epidemiological studies are considered. These include:

- older adults
- people with pre-existing respiratory or cardiovascular disease
- asthmatics
- children
- outdoor workers or people who exercise outdoors
- diabetics
- people in low socioeconomic groups.

Studies conducted in the United States and Canada are given greater weighting because they are thought to be most relevant to the setting of United States air quality standards. Emphasis is placed on studies that have been conducted in numerous cities using standardised methodological approaches, or provide overall effect estimates on combined analysis of information pooled across multiple cities. Studies that provide quantitative effect estimates for populations of interest and

that consider other pollutants in the modelling (multi-pollutant models) are given a higher weighting.

In the United Kingdom, the standards are derived almost entirely from epidemiological data. Toxicological data are used to provide evidence of biological plausibility for an effect observed in the epidemiological studies. The main focus of the hazard assessment is epidemiological studies, both population based and controlled exposure. In cases where robust exposure–response functions are not available from population-based studies, NOAEL or LOAEL data from controlled exposure studies may be used.

The WHO uses epidemiological evidence from human health studies in preference to animal toxicological studies as the basis for their air quality guidelines. Evidence from toxicology and other disciplines is used in support of epidemiological evidence, and in making a judgement about the strength of causality. The results of meta-analyses are used for exposure–response functions, and strict guidelines on selection of studies for the meta-analysis are followed.

The following points highlight the key selection criteria used by the WHO in selecting studies for meta-analysis, as part of the development of air quality guidelines for Europe:

- The number of estimates available for meta-analysis was not a determining factor in selecting studies — that is, there was no compromise on any of the criteria for study selection in order to raise the number of studies included in the analysis.
- The focus of the meta-analysis was on European studies, but if there are not enough studies to perform a meta-analysis, then it may be necessary to consider the inclusion of studies from other parts of the world.
- Only one estimate from each city should be used in the meta-analysis. Where a number of cities had been studied more than once, a mechanism for selecting the appropriate estimate was developed. It was decided to select the latest study that met the selection criteria.
- The initial analysis focused on single-pollutant model results, based on an all-year analysis.
- The 'selected' lag from the time-series studies in the database was used rather than specific lags or combinations of lags.

The WHO defines adverse effects on health on a case-by-case basis, and relies on expert judgement. Usually, there is a range of studies addressing a variety of health endpoints. These endpoints are arranged into a hierarchical structure ranging from severe (mortality) to less severe or mild effects (temporary or reversible physiological or psychological changes).

In the 2000 edition the WHO guidelines for Europe (WHO 2000b) were derived based on either the NOAEL or LOAEL whichever was available. Data from human, animals or *in vitro* studies (plants in relation to ecology) were used. Results that could be put into context with other studies were preferred. All available data were taken into account in an exposure–response curve from which a LOAEL was estimated. Studies conducted in Europe were given preference over studies from other parts of the world, although the international studies were considered as part of an overall weight of evidence for effect.

In setting guidelines on the above basis, uncertainty factors (safety factors) have to be taken into account to provide an adequate margin of safety for sensitive groups within the population. Uncertainty factors were applied to account for extrapolation from:

- animal studies to human studies
- small group to large populations
- general population to sensitive subgroups.

In the 2005 global review of the WHO guidelines (WHO 2006b) there was a move from the threshold concept, with the pollutants being treated as non-threshold pollutants, based on evidence from epidemiological studies because:

- adverse effects can be identified at lower levels through increased sensitivity of study designs
- there could be many thresholds, depending on the health outcome
- there is no single population for whom the threshold fits
- the threshold is always based on uncertainty, because studies have limitations
- lack of evidence does not equate with absence of effect or threshold.

Again, controlled human exposure and toxicological studies were used as supporting evidence in a weight-of-evidence approach.

3.2.6 NHMRC guidance for the NEPC process

In 2006 the NHMRC worked in collaboration with EPHC to develop guidance for the hazard assessment process in standard setting. The resulting document from NHMRC (2006) recommended that standard setting should, if possible, be based on health-effects information from general population exposures. General population studies are preferred because they should inherently include sensitive subpopulations and real-life exposures; however, their usefulness can be limited if there are compromises in exposure estimation and if sensitive subpopulations are not characterised. Also, a negative response in a large population study may mask an effect in a sensitive subpopulation. Thus, the NHMRC (2006) also notes that it may be necessary to give greater weighting to quantitative information from other sources (e.g. controlled clinical studies, animal studies) if the study data are not amenable to quantitative interrogation. Obviously, more weighting is given to animal studies that use the inhalational route of exposure. If such studies are lacking, then it may be necessary to consider extrapolating from ingestion data, but this should be avoided if possible. The NHMRC (2006) recommendations provide the basis for the following steps to assess the validity and usefulness of individual epidemiological studies as part of a weight-of-evidence analysis:

- *Step 1 – Evaluate studies for internal validity*: that is, the adequacy of study design and the extent to which it has validly measured what it intends or purports to measure.
- *Step 2 – Evaluate studies for external validity*: that is, determine whether the results can be validly generalised, extrapolated or transferred to other settings (e.g. climatic, demographic, pollution sources and levels, etc).
- *Step 3 – Evaluate corroboration, contradiction and plausibility:* that is, consider whether the Bradford-Hill view points may be useful here.
- *Step 4 – Make a choice*: that is, select the study or studies that best represent the endpoint of most relevance for setting an air quality standard.

Studies should be selected on the basis of their internal validity (i.e. consideration of features that can eliminate bias, chance and confounding) and external validity (i.e. generalisability). The NHMRC (2006) further recommends that the questions listed below be used to guide the selection of studies for the purpose of air quality standard setting.

3.2.6.1 Controlled human exposure or epidemiological studies

- Was the study an analytic study (cohort or case-control) or an ecologic (time-series study?)
- Were the study population and the comparison population well defined and representative of the general population?
- Was the statistical power of the study sufficient to detect effects if they occurred?
- Did the study address potential confounders or the effects of chance and bias?
- Might biases have been introduced through incomplete follow up?
- Were the health outcomes ascertained using reliable, valid and meaningful measurements?

- Was the exposure assessed using reliable, valid, precise and meaningful measures in relation to the study population?
- Was the relationship between exposure and health outcomes relevant (e.g. averaging time and time sequence)?
- Has misclassification (outcome and exposure) been addressed, quantified or discussed?
- Was the study sufficiently rigorous to support its purported outcomes?
- How has the exposure–response effect been characterised, in terms of linear or monotonic increases in relative risk, threshold, exponential, NOAEL and LOAEL?

3.2.6.2 Animal studies

- Were an appropriate number and spread of doses used?
- Was the maximum tolerated dose (MTD) reached or exceeded?
- Was the validity and purity of the test material specified?
- Were there potentially compromising diseases or pathogens within the test animal groups?
- Were the toxic endpoints well characterised?
- Were the dosing regimens well characterised in terms of stability in the dosing material and measurement of consumption?
- Was pathological assessment undertaken and, if so, was there a transparent method of grading the observed changes?
- Were the record keeping, animal husbandry and clinical observation adequately rigorous, and did the study conform to good laboratory practice or some other form of quality assurance?

Studies that form a reliable basis for standard-setting include the following:

- Well-controlled experimental studies in humans (normal subjects and those with disease predisposing them to increased susceptibility), with single or mixed pollutant exposures (chamber studies), noting that such studies would only be ethically acceptable where a relatively mild adverse effect is monitored (e.g. mucous membrane irritation, mild respiratory impairment or reduction in lung function).
- Well-conducted population-based epidemiological studies.
- Animal studies, where the primary adverse effect is systemic toxicity that may not be readily detected in experimental human or epidemiological studies (e.g. cancer, reproductive or developmental toxicity, or frank neurotoxicity).
- Meta-analyses of observational epidemiological studies (refer to Section 3.4.3).

The selection of studies to be used in a standard setting process will also involve making judgements about:

- the weight of evidence of the available data
- the severity of the effects
- the latency period between exposure and onset of effect.

The NHMRC (2006) conclude that studies in animals should not be the primary basis for setting air quality standards where there are studies in humans (chamber exposures or epidemiological studies) that clearly define relevant exposure–response relationships for the critical endpoint. They can, however, provide information on biological mode of action and biological plausibility for effects seen in epidemiological studies. In using animal studies to complement relevant studies in humans, it is important to understand the relative advantages and limitations inherent in the design of animal studies.

3.2.7 Previous NEPC approaches

The standards in the AAQ NEPM were derived primarily from a qualitative assessment of international epidemiological studies. At the time the NEPM was made, no Australian

epidemiological studies were available to inform the standard-setting process. Results of controlled human exposure studies were used to provide evidence of any thresholds for effects, and toxicological data was used for biological mechanisms and plausibility. The review of the health effects of air pollution used as the basis for the standards raised a number of issues:

- use of controlled human exposure (chamber) studies in determining appropriate ambient air quality levels at which standards could be set
- difficulties in separating the health effects of individual pollutants from the effect of a mixture of the pollutants
- interactions between allergens and pollutants
- complexity associated with determining unambiguous exposure–response relationships.

An additional challenge was the absence of a threshold for adverse health effects for some pollutants — in particular for ozone and PM_{10}.

As previously discussed, the standards for $PM_{2.5}$ were developed through a quantitative risk assessment process. The standards are based primarily on the results of epidemiological studies. Too few local studies were found on which to base a standard. Therefore, exposure–response relationships from international studies were used in the risk assessment. In selecting the studies, a range of criteria was used. These included:

- similar climatic conditions to Australia
- similar demographics to Australia
- effect estimates that were statistically significant
- $PM_{2.5}$ levels were as similar as possible to those experienced in Australia
- similar sources of $PM_{2.5}$
- health effects that had been observed in Australian studies.

The final studies selected had also been identified as key studies in the review of the $PM_{2.5}$ standard in the United States. For both the original NEPM standards and the $PM_{2.5}$ variation, controlled human exposure studies and toxicological studies were used as supporting evidence. Since the development of the standards in the original NEPM and for $PM_{2.5}$, EPHC/NEPC has initiated a number of epidemiological studies in Australia that can be used in setting standards in the future.

3.2.8 Recommended approach to hazard identification

The recommendations for identifying hazards to be considered in setting air quality standards are:

- The hazard assessment should be guided by a scientific and transparent weight-of-evidence approach by an appropriate mix of a team of people with expertise to conduct a weight-of-evidence analysis.
- In selecting studies as part of the analysis, the guiding questions relating to study selection and evaluation (see Section 3.2.6) should be applied. Human-based studies are preferred.
- Epidemiological studies — especially those representative of the general population (including sensitive groups) — should be used, if available. Human-based studies are preferred. Controlled human exposure studies should only be used for short-term health outcomes, when appropriate epidemiological studies are not available.
- Toxicological studies should be used as supporting evidence of biological plausibility and coherence of effect. When no human data are available (as is the case for many air toxics), toxicological evidence can be used for standard setting, with the appropriate application of uncertainty factors.

3.3 IDENTIFICATION OF CRITICAL HEALTH OUTCOMES

The health effects of the criteria pollutants are well documented; they include:

- increases in daily mortality (mainly respiratory and cardiovascular causes)
- increases in hospital admissions and emergency department attendances (respiratory and cardiovascular causes)
- exacerbation of existing disease
- exacerbation of asthma
- increases in medication usage
- low birth-weights and retarded intrauterine growth
- decreases in lung function
- increases in respiratory infections
- irritation of the airways
- changes in heart-rate variability and ability of blood to coagulate.

The effects can be short term (associated with daily changes in air pollution) or long term (associated with annual average levels of air pollution). For air toxics, a wider range of health outcomes have been identified, and the link with causation of disease is sometimes much stronger. For the air toxics, health outcomes include:

- cancer
- reproductive and developmental effects
- central nervous system effects
- respiratory irritation (including eye irritation)
- increases in the incidence of asthma and other chronic lung diseases
- exacerbation of asthma.

In some cases (e.g. carbon monoxide and lead), biomarkers have been used as the basis for setting air quality standards. For carbon monoxide, air quality standards have been based on blood carboxyhaemoglobin levels that have been translated to ambient air concentrations. For lead, blood lead levels linked to intelligence quotient deficits in children have been used. The blood lead levels, as for carbon monoxide, have been translated to an ambient air concentration.

3.3.1 Use of adverse health effects in setting standards

To determine which health outcomes should be the basis for air quality standards it is necessary to consider what constitutes an adverse health effect. For example, are small reversible decrements in lung function an adverse effect?

The American Thoracic Society (ATS) has developed guidelines as to what constitutes an adverse respiratory health effect in relation to air pollution (ATS 2000). The organisation concludes that the following should be considered as adverse effects:

- *quality of life* — measurable negative effects of air pollution on health-related quality of life, whether for individuals with chronic respiratory disease or for the population in general
- *physiological impact* — reversible loss of lung function in combination with the presence of symptoms, or any detectable level of permanent lung function loss attributable to air pollution exposure (small transient loss of lung function by itself should not be considered adverse)
- *symptoms* — air pollution-related symptoms associated with diminished quality of life or change in clinical status
- *clinical outcomes* —detectable effects of air pollution on clinical measures
- *mortality* — any effect on mortality (in interpreting the evidence from short-term time series studies the extent of life-shortening underlying the association needs to be considered)
- *population health versus individual risk* — if the population distribution of exposure shifts towards a higher level, the risk to the population may move to a higher level even though the risk to individuals within the population may remain the same; assuming that the relationships between risk factor and disease are causal, such a shift in the risk profile of the population

should be considered adverse, even in the absence of the immediate occurrence of clinically evidenced illness in individuals.

The NHMRC (2006) refers to the American Thoracic Society (ATS) criteria for determining an adverse health effect in relation to air pollution and thus for their use in setting air quality standards. The US EPA and WHO use similar criteria to identify adverse respiratory health effects. Similar guidelines have not yet been developed for cardiovascular outcomes. However, agencies such as WHO and US EPA use cardiovascular outcomes (e.g. total cardiovascular diseases, ischemic heart disease, dysrhythmias and congestive heart failure) in setting standards.

The US EPA also refers to the ATS statement to determine adverse health effects of air pollution (US EPA 2008). The US EPA notes that in the 2000 update of the ATS statement, there was an increased focus on quality of life measures as indicators of adversity and a more specific consideration of population risk. Exposure to air pollution that increases the risk of an adverse effect to the entire population is viewed as adverse, even though it may not increase the risk of any identifiable individual to an unacceptable level. For example, a population of asthmatics could have a distribution of lung function such that no identifiable single individual has a level associated with significant impairment. Exposure to air pollution could shift the distribution such that no one individual experiences clinically relevant effects; however, this shift toward decreased lung function would be considered adverse because individuals within the population would have diminished reserve function and, therefore, could be at increased risk if affected by another agent.

Air quality standards are usually based on the health outcomes where the strongest weight of evidence of an adverse effect is obtained. In a recent review of air quality guidelines for ozone and particles, WHO (2005b), focused on mortality (all cause, respiratory and cardiovascular), increases in hospital admissions (respiratory and cardiovascular), cough (unspecified, cough with wheeze and tight chest and night cough) and medication usage in symptomatic individuals. Cough and medication usage in symptomatic individuals were used as indicators of worsening of respiratory health, reflecting exacerbation of asthma (WHO 2005b). Evidence for mortality and hospital admissions came from population-based epidemiological studies, and the evidence for cough and asthma from panel studies.

In recent reviews of air quality standards for particles and ozone, the US EPA also focused on evidence from population-based epidemiological studies and panel studies (Abt Associates 2005, US EPA 2005. In the risk assessment conducted for the review of PM_{10} and $PM_{2.5}$ standards, only the most severe and well-understood health outcomes were used. Health outcome were also chosen where the weight of evidence supported a causal relationship or the scientific evidence was sufficiently suggestive of a causal relationship being likely (US EPA 2005).

The final health outcomes used as the basis of the revised standards were mortality (all cause, cardiovascular, respiratory), increases in hospital admissions (respiratory and cardiovascular) and respiratory symptoms not requiring hospitalisations. Decreases in lung function and heart rate variability were not used.

For ozone only mortality (all cause, cardiovascular, respiratory) and increases in hospital admissions (respiratory and cardiovascular) were used for the derivation of possible standards. All other types of studies, including controlled exposure studies, were used in a qualitative manner as part of the hazard assessment, and were part of the evaluation of the weight of evidence for an adverse effect from exposure to ozone.

The NHMRC identifies clinical symptoms related to air pollution as those associated with change in clinical status or diminished quality of life. Air pollution can provoke symptoms in people who are otherwise asymptomatic and apparently healthy, and can trigger symptoms in people with asthma, chronic obstructive pulmonary disease (COPD) or cardiac disease (NHMRC 2006).

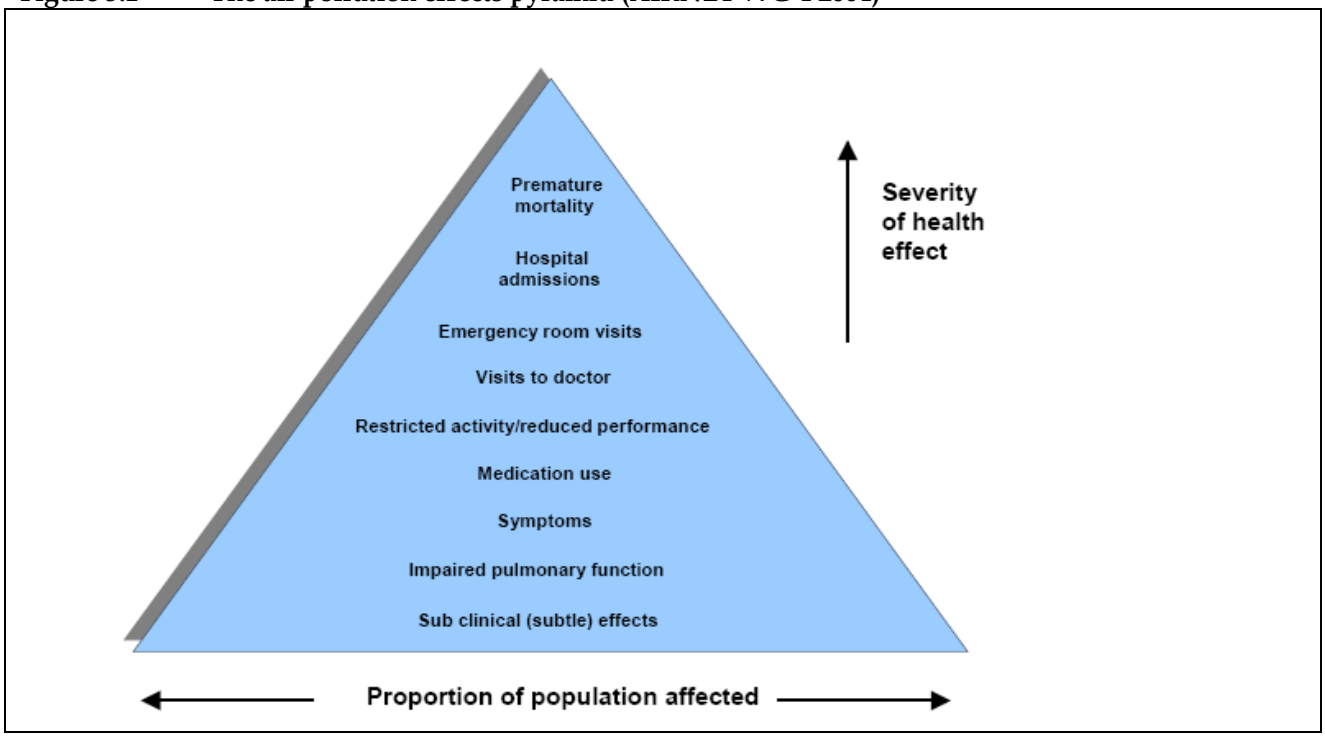
Minor symptoms such as infrequent eye irritation, sneezing or cough should not necessarily be considered adverse. However, wheezing in young adults with asthma is associated with reduced quality of life, and more frequent wheezing, particularly in association with sleep disturbance or symptoms on waking, is associated with severe asthma. Symptom frequency is one component of the classification of asthma severity. A progression of asthma from mild to moderate, or from moderate to severe should be considered an adverse effect (NHMRC 2006).

3.3.2 Health outcomes linked to exposure to air pollution

As discussed above, the health effects linked to air pollution range from severe effects such as mortality to less severe effects such as respiratory symptoms. The less severe symptoms may be experienced by a larger segment of the population, thus posing a greater public health burden than the severe effects. In addition to the direct health effects, a range of flow-on effects have been linked to exposure to air pollution, including loss of work days and school absenteeism.

The range of health effects attributed to air pollution is shown in Figure 3.1 and illustrates the extent to which less critical health effects can affect a wide proportion of the population.

Figure 3.1 The air pollution effects pyramid (AIRNET WG 4 2004)



The NHMRC (2006) state that any increase in mortality related to air pollution is of serious concern.

The ATS (2000) position paper canvassed the possible phenomenon of 'harvesting' — the idea that there is a pool of frail elderly individuals whose death is advanced by a few days following an air pollution event. Careful statistical examination of this concern by Schwartz (2000) suggests that harvesting on this short time scale probably does not occur to any significant extent. It does appear, however, that deaths may be advanced by a few months or more (NHMRC 2006).

When the relationships between risk factor and disease appear to be causal, the ATS considers that a shift in risk-factor distribution, and hence the risk profile of the exposed population, is adverse, even if there is no immediate overt illness.

As levels of exposure to air pollution rise, not only do more people experience adverse effects that are potentially more severe, but individuals who were previously not susceptible may become susceptible to other factors affecting respiratory function (e.g. allergens or viruses), and individuals with sub-clinical disease can become symptomatic. For example, the responsiveness of sensitised asthmatics to an airborne allergen (house dust mite) is increased by immediate previous exposure to nitrogen dioxide (Tunnicliffe et al 1994).

Changes in heart rate and rhythm have been documented following exposure to ozone (Rich et al 2006), fine particles and other vehicle exhaust pollutants (Holguín et al 2003, Dockery et al 2005). However, heart-rate variability is at best a surrogate endpoint, and there is no agreement as to what constitutes a clinically important change. Changes have also been observed in blood coagulability, viscosity and markers of systemic inflammation (Donaldson et al 2000). Although these subclinical effects probably increase the risk of myocardial infarction and sudden cardiac death, and provide evidence for biological plausibility for effect, their use in setting air quality standards is limited to that role.

Research on air toxics has also considered reproductive and developmental effects, reduced performance on neuro-behavioural testing and an increase in neuropsychological disease. Individual papers reporting such effects in humans should be carefully evaluated in the context of comparable data from animal studies.

Air pollution may have measurable negative effects on health-related quality of life, whether for people with chronic respiratory disease or the general population. An increasing body of research on health outcomes has highlighted the importance of health-related quality of life. Quality of life includes the physical domain (e.g. symptoms and exercise capacity), psychological domain and the socioeconomic domain (social function). Validated instruments to measure quality of life have been developed that are sensitive to the impact of environmental factors. These include generic instruments such as the Short Form 36 (SF36) (Ware et al 1992) and disease-specific instruments such as the Asthma Quality of Life Questionnaire (Marks et al 1992), Chronic Respiratory Questionnaire (Guyatt et al 1987) and the St Georges Respiratory Questionnaire (Jones 1993). The minimum clinically important difference has been defined for these instruments (e.g. four points on the St Georges Respiratory Questionnaire). Where the exposure continues over a considerable duration, this concept can be extended to estimate quality-adjusted life years (QALYs).

3.3.3 Sensitive populations

Within the general population, there may be subgroups that are potentially more susceptible to the effects of air pollutants than are others. Human variability (also called intra-species or inter-individual variability) may arise through toxicokinetic or toxicodynamic variability (Dybing and Soderlund 1999), both of which may be due to acquired or inherent factors, or both. A sensitive sub-population is one that demonstrates either of the following (NHMRC 2006):

- an adverse response to an air pollutant occurs at concentrations substantially lower than those that affect most of the population (the concept applies principally to irritant and allergenic compounds)
- the consequences of exposure are more significant than in most of the population (e.g. children may be considered a sensitive population because any irreversible adverse effects may influence their health throughout their life; older adults, especially those with specific co-morbid effects such as cardiac or respiratory failure, may also constitute a sensitive group because the secondary consequences — for example, pneumonia or worsening cardiac failure — may be more serious than in the remainder of the population.

A number of factors can affect the response of an individual to air pollution — genetics, gender, age, pregnancy, health and nutritional status, and lifestyle choices. These factors need to be taken

into consideration when assessing risk to a particular community or population. Sensitive subpopulations may be characterised by:

- clinical history (e.g. of asthma, cardiac failure, chronic bronchitis or cystic fibrosis)
- evidence of airways hyper-responsiveness (e.g. using methacholine or specific challenge tests)
- demographic factors (e.g. age)
- genetic factors (e.g. cystic fibrosis).

Developmental factors — for example, an immature immune system, aging and pregnancy — also affect how an individual may respond to an environmental pollutant. Children have a higher metabolic rate, a more rapid growth rate, a higher percentage of water in their body tissue and an immature immune system that makes them more susceptible to environmental stresses. With aging, kidney and liver function declines, thereby affecting an individual's ability to rid their body of toxicants. In addition, increases in body fat and decreases in body water with age may affect the accumulation of toxicants in the body. During pregnancy, women encounter greater stresses when metabolising and detoxifying, which may influence their response to environmental pollutants and may ultimately affect the foetus. The WHO (2006a) provides detailed guidance on these issues including exposure assessment and developmental stage–specific susceptibilities and outcomes.

Diseases of the kidneys, liver, heart and lungs may predispose an individual to more serious consequences of pollutant exposure than an individual not suffering from such diseases. Cardiovascular and respiratory diseases decrease an individual's ability to withstand environmental stresses, and impaired liver and kidney function affects the body's ability to excrete toxic substance or their metabolites.

Nutritional status is an important factor as nutritional deficiencies can alter the rates of absorption of environmental chemicals, and can induce changes in body composition, which in turn can alter the tissue distribution of chemicals. Dietary factors can strongly influence kidney function and pH of body fluids, and therefore may alter the body's ability to remove toxicants. Low-protein diets and high-carbohydrate diets lead to a decreased rate of detoxification. Similarly, fasting or starving depresses the metabolism and results in reduced clearance of chemicals from the body.

Lifestyle choices such as tobacco smoking, and use of alcohol and other drugs also affect an individual's response to toxicants. Tobacco smoking can act synergistically with several environmental pollutants. Alcohol use may result in brain and liver disorders. As a result, heavy alcohol drinkers who are exposed to environmental pollutants may experience more serious toxic effects than individuals who do not drink.

The groups identified as being particularly susceptible to effects of the criteria air pollutants are:

- older adults
- children
- asthmatics and people with existing respiratory and cardiovascular disease
- diabetics
- foetuses
- low socioeconomic groups.

In addition, people who work or exercise outdoors may be more vulnerable to the effects of air pollution due to increased exposure. Sensitive subgroups have not been well defined for air toxics. Air quality standards are usually set to protect the most sensitive groups within the population. All standards developed in Australia and overseas take this approach. In some cases, however, such as the setting of sulfur dioxide standards to protect asthmatics, it is not possible to protect the most sensitive of these groups.

3.3.4 NHMRC guidance

NHMRC (2006) made recommendations for the selection of critical health outcomes and sensitive populations to be used as the basis for standard setting in Australia. These recommendations can be summarised as follows:

- Any clinically significant effect of air pollution should be considered adverse.
- Mortality is a critical endpoint for setting air quality standards.
- The possibility of sensitive subpopulations within the general population should be considered when setting an air quality standard.
- The impact of air pollution on children's health should be actively considered when setting air quality standards.
- Provided data warrant it, the standard should be adjusted to account for possible increased sensitivity of children.
- In the absence of information showing that children have increased sensitivity to a particular air pollutant, no adjustment for child exposure is needed because the default adjustments for human variability within the adult population adequately protect children.
- Chemicals with a mutagenic mode of action should be evaluated on a case-by-case basis.

The following broad principles should be taken into account in setting air quality standards, to allow for sensitive populations:

- Irritant substances are likely to be of greatest concern to sensitive subpopulations.
- Acceptable levels of such agents should be established, such that they afford substantial protection to sensitive subpopulations.
- The extent of targeted protection must take account of the feasibility of establishing and enforcing specific exposure limits, population-wide consequences (especially the numbers affected and severity of the consequences), and the availability of other management strategies for those affected (e.g. anti-asthma drug therapy).
- No-effect levels in sensitive subpopulations may be best established by controlled exposure studies, but other relevant information may be derived from epidemiological studies. While controlled exposure data for both normal and sensitive subpopulations will often not be available, results from epidemiological studies may be influenced by the presence of members of sensitive subgroups (e.g. asthmatics).
- Safety margins employed in standard setting should be sufficient to include NOAEL in most members of the most sensitive subgroups; typically, these will be individuals with moderate to severe asthma.

3.3.5 Previous NEPC approaches

In the development of the AAQ NEPM standards and the standards for $PM_{2.5}$, mortality and hospital admissions for both respiratory and cardiovascular causes were assessed in a quantitative manner. Standards have been developed to protect the most sensitive individuals including:

- children
- older adults
- people with pre-existing disease (e.g. asthma, COPD or cardiovascular disease).

Particular attention has been paid to the impacts of air pollution on people with asthma, due to the high incidence of asthma experienced in Australia.

The health effects assessed in the development of the $PM_{2.5}$ standard included:

- short-term effects
  - daily mortality (all causes — non-traumatic)
  - daily mortality (respiratory disease)
  - daily mortality (cardiovascular disease)

- daily hospital admissions (asthma)
- daily hospital admissions (cardiovascular disease)
- daily hospital admissions (COPD)

- long-term effects
  - mortality (all cause — non-traumatic)
  - mortality (lung cancer)
  - mortality (cardiopulmonary disease).

3.3.6 Recommended approach for selecting critical health outcomes and sensitive populations

The recommendations for selecting critical health outcomes and sensitive populations as the basis for setting air quality standards are:

- Toxicological studies should be used as supporting evidence of biological plausibility and coherence of effect. When no human data are available (as is the case for many air toxics), toxicological evidence can be used for standard setting, with the appropriate application of uncertainty factors.
- Any clinically significant effect of air pollution should be considered adverse. The NHMRC guidance on determining what constitutes a clinically significant effect, which reflects the ATS position, should be used for identifying adverse health effects.
- Mortality should be considered as a critical health outcome. More sensitive health outcomes (e.g. hospital admissions, emergency department visits, exacerbation of asthma and other respiratory or cardiovascular diseases and reversible decrements in lung function) should also be used where exposure–response data are available. Particular attention should be paid to people with asthma, given the high prevalence of asthma in the Australian population.
- Broader quality-of-life issues, such as work loss days and school absenteeism, should be considered as health outcomes where data are available.
- Sensitive subpopulations within the general population should be taken into consideration; they include children, older adults, people with existing respiratory and cardiovascular disease, asthmatics, diabetics, and low socioeconomic groups.

3.4 DETERMINING EXPOSURE–RESPONSE FUNCTION

One of the Bradford-Hill viewpoints for causation is related to exposure–response or biologic gradient, and refers to the ability to show a relationship between risk factors such as air pollution levels and associated health effects. In essence, if it can be shown that health effects increase with higher exposures, then a positive exposure–response relationship can be stated, which adds strength to a causal interpretation for effects. However, the possibility of confounding by closely related confounders has to be considered. For example, the relationship between nitrogen dioxide and mortality indicates a linear exposure–response relationship, but fine particle exposure is closely related; therefore, the exposure response relationship of nitrogen dioxide has to be evaluated against information from exposure to $PM_{2.5}$.

Most of the health effects of air pollution have been derived from epidemiological studies of human populations in a variety of geographical locations. The benefit of using population-based studies lies in the real-life interpretation of the relationship between these exposures and health outcomes. The levels of air pollutants assessed reflect the exposure of the general population for which the air quality standards are developed, and the observed risks to health at these levels of pollution provide a basis for input into health-based air quality standards.

The results (exposure response or effect estimates) from time-series studies are expressed statistically as risk ratios (relative risk) for a specific health outcome occurring for an incremental change in concentration of the relevant air pollutant (e.g. per 'y' $\mu g/m^3$; y can be set to suit requirements, usually 1 $\mu g/m^3$, 10 $\mu g/m^3$ or the inter-quartile range of the pollutant).

A Poisson regression analysis — the mathematical regression analysis used for small counts of events in a large population — is used when calculating the relative risk based on daily numbers of adverse health events. Resulting effect estimates are based on the underlying health effects measured over the whole concentration range experienced by the study population. The relationship between exposure and response in a regression model can be visualised as a linear curve, with the steepness of the slope reflecting the likely response between concentration and health outcome, including confounders. Other concentration response functions derived from panel studies are based on logistic regression, giving rise to odds ratios (WHO 2005b).

As a rule, relative risk derived from a time-series or cohort study is expressed as a decimal, where a value of 1.00 means no difference in risk (e.g. as many hospital admissions at the higher exposure as at the lower exposure). A relative risk of 2.00 means twice the risk or, when expressed as percentage risk, a 100% increase, while a risk ratio of 1.006 can be also expressed as an increase in risk of 0.6%. If this risk has been calculated for a change of 10 $\mu g/m^3$ of exposure (e.g. in PM_{10}), it expresses a 0.6% increase of risk for every increment of 10 $\mu g/m^3$.

In setting standards, exposure–response functions are combined with baseline health data for a given population (e.g. Melbourne) and exposure data to provide an estimate of risk (e.g. number of deaths or hospital admissions). These risk estimates are calculated for a range of health outcomes and sensitive groups, because the exposure–response functions vary with the health outcome being assessed and the population group to which they apply (e.g. the whole population, or sensitive subgroups such as children or older adults). Exposure–response functions can be obtained from sources such as:

- epidemiological studies conducted in individual cities (single city studies)
- epidemiological studies conducted using the same statistical methodology in multiple cities (multi-city studies)
- meta-analysis that combines the effect estimates obtained from single city studies

The use of exposure–response data from these types of studies is discussed further below.

3.4.1 Single-city studies

Epidemiological studies have been conducted to investigate the impacts of air pollution in individual cities; the exposure–response functions derived from these studies represent the response of the population of an entire city. In many cases, these studies are independent of those in other cities; e.g. they may use different statistical approaches and different ways to adjust for potential confounders, such as weather or episodes of increased influenza illness. High-quality data from a specific city can be used to establish 'city-specific' risk estimates (NHMRC 2006).

In the United States, single-city exposure–response functions are used as the basis for risk assessment in developing NAAQS. Single-city estimates are combined with baseline health data for a range of health outcomes and sensitive populations, and with exposure data for the city for which the exposure response function has been derived. Risk is assessed in a number of United States cities, using data specific to those cities, to address heterogeneity in the results of epidemiological studies investigating the impact of air pollution on health conducted in various parts of the United States.

3.4.2 Multi-city studies

Multi-city studies have been conducted to strengthen findings and to achieve a national exposure–response function. These studies combine time-series data from individual cities that have used the same study protocol, to obtain an overall exposure–response function for specific health outcomes that represent the response of most of the population of interest. Heterogeneity in the results of the studies is visualised using forest plots and other statistical tests. Multi-city studies reduce the effect of bias potentially present in other studies either through post-study determination of models that

support positive results or through underreporting of studies with negative results (publication bias[2]) (WHO 2004a).

Multi-city studies include the National Morbidity and Mortality Air Pollution (NMMAP) study in the United States, which involved 100 cities; the Air Pollution and Health: European Approach (APHEA) 1 and 2 studies in Europe; the Air Pollution and Health: Europe and North America (APHENA) study, which combines the results of studies from cities in the United States, Canada and Europe; and the Public Health and Air Pollution in Asia (PAPA) studies. In Australia, a multi-city study was conducted in four cities (Melbourne, Sydney, Brisbane and Perth) (Simpson et al 2005ab). EPHC commissioned an expansion of this study (i.e. with additional cities and pollutants), to inform the review of the standards in the AAQ NEPM. The expanded study examined the associations between air pollution and mortality and hospital admissions in seven Australian and two New Zealand cities (Barnett et al 2005, 2006). This study combined the results of single cities into a random effects meta-analysis to obtain exposure–response functions for specified health outcomes.

3.4.3 Use of meta-analyses in air quality standard setting

The third method for obtaining exposure response functions is by meta-analysis — a technique that can be used to combine the effect estimates obtained from single-city studies. The aims and statistical techniques are again to provide a more precise estimate of the exposure–response function. The validity of a meta-analysis depends on the quality of the systematic review of the underlying studies, and the review should include all relevant studies. A comparison of results from the NMMAP study multi-city analysis, which was based on 95 of the 100 cities, with a meta-analysis based on 39 single cities in relation to acute ozone mortality, indicated that both designs showed strong linear associations with increasing ozone concentrations. The meta-analysis of the single city studies showed higher effect estimates, indicating possible publication bias (i.e. the tendency to only publish positive studies) or other methodological irregularities (Bell et al 2005). The authors concluded that it was important to assess these potential biases when meta-analytic approaches are used to determine exposure–response functions.

When conducting meta-analyses of multi-city and single-city studies, heterogeneity between studies must be considered. This technical term refers to differences or variability in study results; for example, exposure–response estimates in the studies, which may occur, for example, due to differences in study methodologies, populations and exposure levels present in the studies to be combined. Heterogeneity should be investigated (by statistical methods and subgroup analysis) before deciding which studies to include in the review. Single-city studies included in a meta-analysis do not necessarily share the same a priori study protocol as is the case with multi-city studies. The WHO meta-analysis, for example, obtained all time-series and panel studies available for European cities at the time, including multi-city studies (WHO 2004a).

Overall, meta-analysis and multi-city studies are important for standard setting because they provide information across the country, and potentially provide more precise exposure response estimates. They also offer insight into potential heterogeneity between cities or geographical regions. Finally, the meta-analytic database can be used for sensitivity analysis of the effect estimates in relation to inclusion and exclusion of specific cities or regions, or sensitive subgroups, and in relation to exploring the curve behaviour, by truncation at the upper or lower end of exposure.

Guidance on the interpretation of meta-analyses in environmental health risk assessment has been provided by the WHO (2000a) and is summarised below:

[2] Publication bias refers to the fact that positive results have better chances of being published, are published earlier, and are published in journals with higher impact factors. Conclusions exclusively based on published studies may be misleading.

- the meta-analysis should have a protocol that specifies the objectives and methods
- inclusion criteria for studies should be broad rather than narrow
- characteristics of primary studies should be assessed qualitatively rather than using a global quality score
- meta-analysis can be performed by inverse variance weighting or random effects models
- the effects of publication bias should be assessed by sensitivity analysis (graphical techniques such as a funnel plot[3] may also be appropriate)
- overall heterogeneity should be assessed because this may identify susceptible groups, exposure conditions, etc
- meta-analytical methods that may be used to compare studies include stratified analysis and meta-regression
- sensitivity analysis should be performed to assess the robustness of summary estimates to the inclusion and exclusion of particular studies
- quantitative summary estimates provide useful input to health impact assessment.

In the development of the 2005 global WHO air quality guidelines, the results of meta-analyses were used as the basis for the guidelines.

3.4.4 Uncertainties and limitations of exposure response estimates

This section discusses issues that can affect the accuracy of the exposure–response estimates obtained through either single-city or multi-city studies or through meta-analyses. These issues must be taken into account when selecting exposure–response functions for setting air quality standards.

Uncertainties in relation to exposure–response estimates are different from those drawn from the weight-of-evidence analysis (bias, confounding, chance findings and study design) that evaluates the relevance of the association between air pollution and health effects. Effect estimates can change in size depending on various factors, but can still support the overall evidence.

The WHO identifies the following uncertainties that need to be taken into account in evaluating exposure response functions for setting air quality standards (WHO 2006b):

- publication bias may increase the summary estimates in meta-analyses
- estimates may be biased due to post hoc analysis
- limited exposure ranges may limit the ability to detect real threshold effects and may influence the shape of the response curve
- exposure influence estimates may be misclassified by shifting the estimate towards the null-hypothesis
- co-pollutants or other unidentified factors may confound the results.

Positive results bias is a type of publication bias that occur when authors are more likely to submit, or editors to accept, positive rather than null (negative or inconclusive) results. In contrast, 'the file drawer problem' is a tendency for negative or inconclusive results to remain hidden and unpublished. Even a small number of studies lost in the file-drawer can result in a significant bias.

Outcome reporting bias occurs when several outcomes within a trial are measured, but are reported selectively depending on the strength and direction of those results. This is seen to be hypothesizing after the results are known.

[3] A funnel plot traditionally has a measure of study size on the vertical axis as a function of effect size on the horizontal axis. Large studies appear toward the top of the graph, and tend to cluster near the mean effect size, whereas small studies appear toward the bottom of the graph and are dispersed across a range of values. This is the result of more sampling variation in effect size estimates in the smaller studies. The pattern tends to resemble a funnel.

Due to the many types of bias, published studies may not be truly representative of all valid studies undertaken. In turn, this may distort meta-analyses and systematic reviews of studies on which health-based air quality standards increasingly rely. When undertaking meta-analyses and systematic reviews, publication bias should be taken into account when identifying the studies to include in the review. A thorough search of unpublished studies should be performed, to minimise the effects of publication bias. The use of analytical tools such as a funnel plot can also quantify the effects of publication bias.

In the absence of publication bias, the studies in a funnel plot would be distributed fairly symmetrically around the combined effect size. However, if there is publication bias, the bottom of the plot would show a higher concentration of studies on one side of the mean than on the other. This would reflect the fact that smaller studies are more likely to be published if they have larger than average effects, which makes them more likely to meet the criterion for statistical significance.

3.4.5 Threshold and non-threshold issues for setting air quality standards

There are two approaches to the assessment of exposure–response functions — threshold and non-threshold (explained below). Essentially, these two approaches depend on assumptions about the shape of the exposure–response curve at low exposure levels. A threshold approach is applied to many air toxics. Current epidemiological evidence is that the non-threshold approach applies to all current criteria pollutants, and possibly to genotoxic air toxic compounds. The choice of exposure–response approach is taken on a case-by-case basis.

3.4.5.1 Threshold concept

The health data that feed into the risk assessment process come from a variety of studies, including epidemiological and toxicological research. In the threshold approach, a threshold level is considered to exist, below which exposure to a pollutant is without adverse effect. In mechanistic terms, a threshold is possible because of biological mechanisms that act to metabolise or excrete a toxin, or repair any damage up to a certain dose.

Thresholds for adverse health effects are found for some carcinogenic air pollutants and many air toxics. For these pollutants, threshold approaches identify either a NOAEL or, if the NOAEL cannot be determined, the LOAEL from key (critical) human or animal studies. If only a LOAEL is available, a NOAEL can be derived by applying an uncertainty factor.

For air pollutants where NOAELs or LOAELs are derived from controlled exposure studies, there is uncertainty as to whether the NOAEL or LOAEL applies to populations other than those included in the study, or at concentrations outside the range of exposure used in the study. Controlled exposure studies are limited to generally healthy adult males, and the most sensitive groups within the population are not included. Uncertainty may also result from selection — for example, depending on whether the NOAEL is a true NOAEL or whether a lack of effect is possible at higher exposures, and similarly, where the LOAEL actually sits on the exposure response curve.

Several types of extrapolations may be necessary to assess risks of toxicity in human populations from animal studies. The 'uncertainty' or 'safety' factors are applied to the NOAEL or LOAEL to derive a standard that will protect sensitive members of the population. These factors are introduced due to the need (for example) to extrapolate from animal data, and to allow for the small sample sizes in the underlying toxicological and clinical studies (human variability).

Where threshold approaches are used in setting air quality standards, the issue of the selection of uncertainty factors becomes critical. Considerable guidance on the use of uncertainty factors is available, and the use of the approach by the International Programme on Chemical Safety (IPCS, 2005) is recommended by the NHMRC (2006).

There is no explicit use of the exposure–response curve in the NOAEL/LOAEL approach, no estimate of risk at the NOAEL, and, no extrapolation to lower risks at lower doses. An alternative approach is to use a benchmark concentration (BMC) corresponding to a specified low level (1–10%) of risk as the point of departure. A lower confidence limit for the BMC is used to allow for experimental variation, and an additional uncertainty factor is introduced to allow for the point of departure being an actual level of risk. This approach is further discussed in the NHMRC hazard assessment guidance (NHMRC 2006).

3.4.5.2 Non-threshold concept

For the air pollutants ozone and particles there is considered to be no threshold for effect because no clear threshold has been able to be determined from epidemiological studies. More recently, criteria pollutants previously thought to be threshold pollutants (e.g. nitrogen dioxide and carbon monoxide), are also now being treated as non-threshold pollutants. Studies into health effects of particulate matter and ozone conducted in 23 cities in Europe and in 98 cities in the United States observed short-term health effects over a large exposure range. Similar results have been found in Australian studies (Barnett et al 2005, 2006; Simpson et al 2005ab). Linearity of exposure–response across observed exposure levels has also been reported from long-term follow up of mortality in association with fine particle levels (Pope et al 2002).

The presence of linearity and absence of clear thresholds is most likely due to observation of large populations that, unlike small experimental groups, include susceptible sub-populations and individuals. Variability in susceptible factors in large populations is inevitably larger than in small experimental groups; thus, threshold observations may be elusive or irrelevant in such a context. Findings of thresholds in experimental studies conducted with a relatively uniform population (which may not include sensitive populations) therefore do not contradict linear concentration-response findings from large population studies. The WHO adopted the non-threshold approach in developing the 2005 air quality guidelines (WHO 2006b).

Non-threshold approaches are also applied to certain toxicological endpoints (e.g. carcinogenicity) where a threshold is assumed not to exist. This assumption has its origins from an early premise by Ames et al (1973) that all carcinogens are mutagens, and that a single mutation or DNA damage event was sufficient to initiate a process leading to cancer. More recently, it has been found that not all carcinogens are DNA reactive (i.e. they do not have direct mutagenic activity); therefore, the dose–response curve is assumed to be threshold for non-genotoxic carcinogens and non-threshold for genotoxic carcinogens.

'No threshold' implies that health effects are expected at all pollutant concentrations. Therefore, it is important to quantify the level of risk that is inherent in any standard for a non-threshold pollutant. For setting standards, the 'no-threshold' concept requires knowledge of the shape of the concentration–response curve, to calculate public health benefits. Because this approach provides estimates of risk at all doses, it allows computation of comparative risk in the sub-experimental range (compared to threshold approaches, which require an experimentally produced NOAEL or LOAEL), comparison of potency between chemical agents at a particular risk level, and estimation of the increased risk if particular doses are exceeded.

A disadvantage of non-threshold approaches is that the risks extrapolated at doses below the experimental range can vary considerably, depending on the mathematical regression model chosen. Also, the numerical expression of the estimated level of risk falsely gives an impression of precision that tends to mask the assumptions required for the calculation.

3.4.6 Use of overseas exposure–response data in Australia

The use of overseas exposure response functions was the subject of significant debate during the development of the AAQ NEPM standards. Concern was raised that the underlying health status of the population in overseas cities may not be relevant in Australia (this is particularly important for asthma, given that Australia has one of the highest rates of asthma in the world) and that air pollution levels in Australia are generally lower than cities in North America and Europe, where most of the epidemiological studies had been conducted. Since that time, several studies conducted in Australia have shown that the effects observed overseas are also observed here, although the exposure–response data may differ.

Although local data need to be available to support the setting of air quality standards for Australia, there will always be some reliance on overseas data, given the extensive databases available. The RATF report (NEPC 2000) discusses the importance of the Australian database in validating the transferability of overseas data to the Australian situation. So far, the findings of Australian studies are consistent with those observed in studies overseas.

The NEPC RATF and RAWG were tasked respectively with identifying risk assessment approaches that could be used in the development of air quality standards in Australia and the data required for such risk assessments to be undertaken. Both groups concluded that key health research should be undertaken in Australia to inform the development of air quality standards, but that overseas epidemiological data would always be needed, due mainly to constraints on resources available for research in Australia and the relatively small size of the Australian population, both of which limit the types of studies that can be conducted.

Overseas epidemiological studies were used in the development of the standards for $PM_{2.5,}$ being used to identify key health endpoints and exposure–response relationships. The results of the Australian studies were used to support the overseas data. In particular, in choosing the appropriate health endpoints for the basis for the standards, only those for which associations between $PM_{2.5}$ had been observed in Australian studies were used. This limited the range of health outcomes considered.

One limitation of using overseas studies was that exposure–response relationships were not always available for the key health outcomes of importance in Australia. For example, children with asthma were a sensitive group identified in the Australian population, and Australian air pollution epidemiological studies had found strong links between $PM_{2.5}$ and hospital admissions for asthma in this group; however, overseas exposure–response data were not available for this age group. As a compromise, asthma admissions were considered for all ages, instead of for the most sensitive group. The resulting standards are considered to be protective of children because the levels at which they have been set are lower than those at which health effects have been previously observed. As $PM_{2.5}$ is a non-threshold pollutant, any level set as an air quality standard has an inherent risk associated with it — the aim is to minimise that risk as far as practicable.

In conducting risk assessments and setting air quality standards, Australian studies would provide information on whether overseas data realistically estimates potential health risk to Australians. In assessing the transferability of overseas exposure response relationships to Australia the following issues need to be considered:

- Were the studies conducted in cities that experience similar climatic conditions to Australia?
- Are the demographics similar to those in Australian cities?
- Are the air pollution levels similar to those experienced in Australia?
- Are the sources overseas and in Australia similar for each pollutant?

- Have the health effects identified been observed in Australian studies, or is there strong evidence to suggest that such effects would be expected to be observed in Australian cities?

Where well-conducted epidemiological studies or meta-analyses have been undertaken in Australia, the exposure–response function from these studies should be used in preference to overseas studies. Also, if there is strong evidence from overseas studies for an effect (e.g. long-term impacts of particles on mortality) that has not been investigated in Australia, this should not be a reason to exclude this health outcome from being assessed as part of the development of the standards. However, the uncertainties involved in using international exposure–response data need to be clearly identified and documented in the risk characterisation.

3.4.7 NHMRC guidance on determining exposure response

The NHMRC (2006) recommend that meta-analyses of high-quality observational studies are preferable for setting standards in Australia. Such meta-analyses are relatively uncommon. More reliable data are likely to be derived from a qualitative review of high-quality 'pivotal' studies. High-quality Australian data should be included amongst this analysis.

The NHMRC further recommend that exposure–response data from a specific city might be used to establish 'city-specific' risk estimates if they are of sufficiently high quality. As stated above, data of limited quality rarely take precedence over well-conducted studies from elsewhere. Extrapolation of data from elsewhere requires that the nature of the exposure and characteristic of the population be similar. Pooled estimates are only reliable if there is homogeneity among the individual studies.

The NHMRC (2006) recommend that, where the key health effects are direct and of a relatively mild nature, experimental studies with humans (controlled chamber exposures) are likely to provide data relevant to standard setting. However, where the health effects are severe or of a delayed onset, and experimental exposure of humans is unethical, well-conducted epidemiological studies, supported by valid meta-analyses or animal experimental studies (or both), are more likely to be useful.

Meta-analysis should be conducted only when original studies are of similar design, use comparable populations, and measure exposure and outcome in similar ways. Appropriate statistical methods should be used to assess meta-analyses, including a test for heterogeneity. Pooled results should not be relied upon for standard setting in the presence of significant heterogeneity between the studies.

3.4.8 Previous NEPC approaches to determining exposure response

As discussed previously the development of the standards in the AAQ NEPM and the subsequent standard for $PM_{2.5}$ relied heavily on overseas exposure–response data. The results from studies conducted in Australia were used as supporting evidence.

3.4.9 Recommended approach for determining exposure response

The recommendations for determining exposure response as the basis for setting air quality standards are:

- A non-threshold approach should be taken for the current criteria pollutants and genotoxic carcinogens. For the criteria pollutants, this approach is based on the evidence from epidemiological studies that indicate that there is no threshold for effects. A threshold approach can be adopted for non-genotoxic carcinogens and other air toxics.
- Exposure–response data obtained from meta-analyses or multi-city studies should be used, provided the primary data are homogeneous. If significant heterogeneity is present, then

exposure–response data for single cities should be used, and risk assessed for the relevant cities for which data are available.

- Exposure–response data from well-conducted Australian studies are preferred. If high-quality Australian data for a particular health outcome are not available, then exposure–response data from overseas studies should be used provided the nature of the exposure and demographics of the population are similar. If overseas data are used, then the uncertainties associated with the use of the data need to be well documented.

3.5 THE APPLICATION OF THE PRECAUTIONARY PRINCIPLE AND ENVIRONMENTAL EQUITY

To meet the requirements of the NEPC Act in developing NEPMs, including air quality standards, a number of policy issues need to be considered. Air quality standards are not simply based on the protection of health, although this is the major consideration, but will also incorporate consideration of the key issues, such as:

- application of the precautionary principle
- environmental equity and justice issues
- costs and benefits associated with the standards.

These issues will be taken into account in determining the form and numerical value of a standard for a particular pollutant. These considerations differentiate an air quality standard from a guideline value.

In terms of hazard assessment the application of the precautionary principle and environmental equity issues focus on identification of sensitive subgroups and the use of international data where data from Australian studies are not sufficient to use as the basis of air quality standards.

3.5.1 Requirements of the NEPC Act

The general purpose of NEPMs, as specified in the NEPC Act, is to ensure that:

- people enjoy the benefits of equivalent protection from air, water and soil pollution, and noise, wherever they live
- decisions taken by business are not distorted, and markets are not fragmented by variations between jurisdictions in relation to the adoption or implementation of major environment protection measures.

The NEPC Act also requires that, when making a NEPM, potential economic and social impacts are considered, and that administration of the NEPM is simple, efficient and effective.

3.5.2 Precautionary principle

The IGAE requires that consideration be given to, amongst other things, the precautionary principle when developing environmental policy in Australia. The IGAE defines the precautionary principle as follows (IGAE 1992):

> 'Where there are threats of serious or irreversible environmental 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 careful evaluation to avoid, where practicable, serious or irreversible damage to the environment, and an assessment of the risk-weighted consequences of various options.'

In recent years, there has been considerable debate internationally about the use of the precautionary principle in the development of air quality standards. The principle is often used as a tool for decision making in situations of scientific uncertainty, and there is growing recognition that current scientific knowledge is often inadequate for characterising the risks associated with

exposure to environmental hazards. (The approach to the application of the precautionary principle overseas is summarised in Appendix 3)

In the development of the standards contained in the AAQ NEPM, including the advisory reporting standards for $PM_{2.5}$, the precautionary principle has been applied in determining the health effects of the pollutants under consideration, dose–response relationships, appropriate uncertainty factors and identification of knowledge gaps. The EPHC has conducted research specifically aimed at generating information to fill those gaps; this research includes the multi-city mortality and morbidity study, and the children's air pollution and health study.

The precautionary principle was also applied in setting air quality standards based on international data, when no relevant Australian studies were available. On the strength of evidence of effect from international studies, it was considered that a precautionary approach would be to set standards based on the international data, and to generate local information for the review of the standards.

The key issue in applying the precautionary principle to setting standards is the consideration of vulnerable or disadvantaged groups, such as children or older adults. A significant portion of the burden of disease from air pollution falls on older adults. Australia's population is ageing, and life expectancy is predicted to increase significantly over the next few decades. A precautionary approach to setting air quality standards should take into account likely impacts on this segment of the population, to ensure that quality of life is not affected.

The effect of air pollution on older adults and children is discussed in detail in Section 3.3. In addition to these sensitive groups, Indigenous Australians may also be more vulnerable to the effects of air pollution, due to existing health conditions and socioeconomic status — factors that have been identified as making people more sensitive to exposure to air pollution. This is discussed further in Section 3.5.3.

3.5.3 Environmental and social equity issues

Environmental and social equity issues are closely entwined with taking a precautionary approach. Not only are specific segments of the population more susceptible to the effects of air pollution due to pre-existing disease or age, but certain segments of the community are exposed to a greater risk from exposure to air pollution due to low socioeconomic status or ethnicity.

International studies have shown that people in low socioeconomic groups are more vulnerable to the effects of air pollution because, for example, they:

- can have a poorer health status, and have greater difficulty in accessing medical care
- tend to live in areas with higher pollution levels (e.g. next to major roads or industrial facilities, as land and house prices are usually lower)
- tend to live in housing that is of older stock; such housing tends to allow greater infiltration of outdoor air, thus increasing exposure to air pollution
- may have limited access to health-support services and have a poorer than average education; hence, they are less likely to seek assistance.

As is the case for increased susceptibility due to pollutants and co-exposure to pollutants, the socially disadvantaged are not usually specifically taken into consideration when establishing an air quality standard. However, many epidemiological studies stratify the exposed population according to socioeconomic status, study these factors or attempt to control for them. When considering human studies, it is important to determine whether the dose-response data are for, or inclusive of, vulnerable populations. The supporting documentation for the standard should specifically address this issue (NHMRC 2006).

Air quality monitoring is not usually targeted towards low socioeconomic areas because the focus is on whole populations. Therefore, the levels of pollution that these communities are exposed to, and their subsequent risk, are often unknown. To ensure that standards meet the requirement of the NEPC Act of providing equivalent protection for all Australians, issues faced by these vulnerable groups should be considered when developing air quality standards.

A further consideration for Australia is the impact of air pollution on Indigenous communities that fall into the low socioeconomic grouping. Although many such communities live outside urban areas they can still be significantly affected by air pollution, in particular by particles from fires and dust, in remote areas of Australia.

Information is generally limited on the exposure to air pollution and the subsequent health risk for low socioeconomic groups in urban areas and for Indigenous populations in nonurban areas. A recent exposure–response study in Darwin on the health effects of PM_{10} from ambient biomass smoke clearly indicated a disproportionate risk for respiratory and cardiovascular hospital admissions in the Indigenous population (Johnston et al 2007). Consideration of the precautionary principle and environmental equity means that these groups should be considered in the setting of air quality standards. However, specific studies to establish the effects of air pollution on Indigenous Australians are limited and do not provide sufficient information to derive exposure-response relationships for assessing the risk posed to this population when developing air quality standards.

Indigenous populations suffer more ill health than other Australians (AIHW 2006). Life expectancy is considerably lower for Indigenous populations that for non-Indigenous populations — 59 years for Indigenous males compared with 77 years for non-Indigenous males, and 65 years for Indigenous females compared with 82 years for non-Indigenous females (AIHW 2006). About 75 per cent of Indigenous populations live in urban areas in Australia and are vulnerable to the impacts of urban air pollution (ABS 2003).

ABS (2004) reports indicate that Indigenous Australians are disadvantaged across a range of socioeconomic factors that affect health including:

- lower incomes
- higher rates of unemployment
- poorer education achievements
- lower rates of home ownership.

As discussed earlier, a number of existing health conditions make people more vulnerable to the effects of air pollution; for example:

- respiratory disease (especially COPD, asthma, bronchitis and emphysema),
- cardiovascular disease
- diabetes.

Low birth weights have also been linked with exposure to air pollution.

According to the AIHW (AIHW 2006), Indigenous Australians have higher rates of all these conditions compared with non-Indigenous populations. Table 3.3 summarises some of the key health outcomes of concern with respect to air pollution and comparison between Indigenous and non-Indigenous populations.

Table 3.3 Prevalence of health conditions: Indigenous and non-Indigenous people, 2004–2005, (as percentages and adapted from AIHW 2006)

Health condition	Indigenous	Non-Indigenous
Respiratory diseases	30	29
Asthma	16	10
Circulatory problems and diseases	22	17
Diabetes mellitus	12	4
Low birth-weight	13	6

The EPHC Cooperative Studies Working Group recommended that research be undertaken to investigate the impact of air pollution on Indigenous populations as a priority. Further consultation with the Indigenous Health Forum identified that there were more significant environmental factors affecting Indigenous health that would take priority over an air pollution study (e.g. sanitation, overcrowding and clean water supplies). Until research is conducted, Indigenous populations cannot be explicitly considered in setting air quality standards; however, the margin of safety included in the assessment of other sensitive subpopulations (e.g. children, older adults and people with existing disease) should provide protection for this population.

International approaches to environmental equity are summarised in Appendix 4.

3.5.4 Recommended approach for consideration of the precautionary principle

The recommendations for considering the precautionary principle as part of the process of setting air quality standards in Australia are:

- The WHO framework for the application of the precautionary principle should be used as a guide.
- Where exposure-response functions are not available in Australia for particular health outcomes, and international studies provide strong evidence of an effect or existence of a susceptible group, then international data should be used.
- Children and older adults should be specifically considered, to ensure that quality of life, now and in the future, is not adversely affected by air pollution.

3.5.5 Recommended approach for dealing with environmental equity issues in Australia

The recommendations for dealing with environmental equity issues as part of the process of setting air quality standards in Australia are:

- Where possible, exposure–response data for affected groups should be used in the risk assessment process to address environmental equity issues.
- All sectors of the Australian population, including vulnerable groups such as low socioeconomic groups and indigenous populations, should be identified and provided the opportunity to be consulted regarding setting air quality standards.

3.6 ESTABLISHING A HAZARD ASSESSMENT EXPERT GROUP

The complexity and specialist knowledge required to undertake the hazard assessment may require skill sets broader than those available within the jurisdictions' environment or health sectors. This will require the NEPC to consider the establishment of a project-based air quality hazard assessment expert group to provide advice and guidance on aspects of health hazard assessment, in developing NEPMs, NEPM reviews and NEPM variations. Appendix 1 provides suggested terms of reference for consideration should NEPC decide to establish an expert panel to assist with the hazard assessment component of standard setting.

3.7 RECOMMENDED APPROACH TO HAZARD ASSESSMENT

The recommendations for hazard assessment as the basis for setting air quality standards are:

- The hazard assessment should be guided by a scientific and transparent weight-of-evidence approach by an appropriate mix of a team of people with expertise to conduct a weight-of-evidence analysis.
- In selecting studies as part of the analysis, the guiding questions relating to study selection and evaluation (see Section 3.2.6) should be applied. Human-based studies are preferred.
- Epidemiological studies — especially those representative of the general population (including sensitive groups) — should be used, if available. Human-based studies are preferred. Controlled human exposure studies should only be used for short-term health outcomes, when appropriate epidemiological studies are not available.
- Toxicological studies should be used as supporting evidence of biological plausibility and coherence of effect. When no human data are available (as is the case for many air toxics), toxicological evidence can be used for standard setting, with the appropriate application of uncertainty factors.
- Any clinically significant effect of air pollution should be considered adverse. The NHMRC guidance on determining what constitutes a clinically significant effect should be used for identifying adverse health effects.
- Mortality should be considered as a critical health outcome. More sensitive health outcomes (e.g. hospital admissions, emergency department visits, exacerbation of asthma and other respiratory or cardiovascular diseases and reversible decrements in lung function) should also be used where exposure–response data are available. Particular attention should be paid to people with asthma, given the high prevalence of asthma in the Australian population.
- Broader quality-of-life issues, such as work loss days and school absenteeism, should be considered as health outcomes where data are available.
- Sensitive subpopulations within the general population should be taken into consideration; they include children, older adults, people with existing respiratory and cardiovascular disease, asthmatics, diabetics, and low socioeconomic groups.
- A non-threshold approach should be taken for the current criteria pollutants and genotoxic carcinogens. For the criteria pollutants, this approach is based on the evidence from epidemiological studies that indicate that there is no threshold for effects. A threshold approach can be adopted for non-genotoxic carcinogens and other air toxics.
- Exposure–response data obtained from meta-analyses or multi-city studies should be used, provided the primary data are homogeneous. If significant heterogeneity is present, then exposure–response data for single cities should be used, and risk assessed for the relevant cities for which data are available.
- Exposure–response data from well-conducted Australian studies are preferred. If high-quality Australian data for a particular health outcome are not available, then exposure–response data from overseas studies should be used provided the nature of the exposure and demographics of the population are similar. If overseas data are used, then the uncertainties associated with the use of the data need to be well documented.

4 EXPOSURE ASSESSMENT – PRINCIPLES AND GUIDANCE

This section discussed the role of exposure assessment in setting ambient air quality standards, and examines Australian and overseas approaches to exposure assessment. The section also provides advice on the adequacy of air monitoring data in Australia to conduct exposure assessment as part of the risk assessment process. The following sections draw heavily on the findings of the RAWG, established by NEPC to build on the work of the RATF.

4.1 ROLE OF EXPOSURE ASSESSMENT IN STANDARD SETTING

Exposure of humans to air pollution can be defined as 'the event when a person comes into contact with a pollutant of a certain concentration during a certain period of time' (WHO 2006). Exposure occurs along the 'environmental pathway' between concentration and dose; this pathway can be expressed as follows:

source → emissions → concentrations → exposure → dose → health effects

Exposure differs from concentration, which is a quantitative measure of the amount of pollutants within an airshed. High concentrations of air pollutants do not necessarily result in high exposures. For example, air pollution concentrations may be high near a source but high exposures will occur only if people spend time near that source.

Exposure also differs from dose, which is the amount of pollutant that actually enters the body. The dose will be defined by the exposure, by factors specific to the pollutant (e.g. its solubility or deposition in the lung) and by physiological factors such as the person's level of activity.

In setting air quality standards, exposure assessment is critical to the risk assessment process. The quality of the exposure assessment will depend on the available air monitoring data and the method used to assess population exposure. Such approaches range from simply averaging the pollutant levels found at air monitoring stations through to complex computer models that require significant monitoring data and modelling. The particular overall approach used needs to be decided on a pollutant-by-pollutant basis, taking into account the available air monitoring data and the types of epidemiological studies used to determine exposure–response relationships (RAWG 2002).

Most research on the health effects of air pollution has focused on respiratory and cardiovascular effects that occur following inhalation of the pollutant. Exposure assessment for air quality management and health risk assessment requires exposure estimates that are accurate, biologically relevant, cover the critical exposure period and quantify the range of exposure levels within the population under study (WHO 2006).

Most studies assessing the health effects of air pollution have used ambient air monitoring data as a proxy for exposure to air pollution, because this is relatively simple and convenient to measure. Air quality standards and guidelines focus primarily on ambient (i.e. outdoor) air quality.

'Total personal exposure' refers to the total exposure people experience in all microenvironments (e.g. outdoors, indoors and in vehicles) where they spend their time (WHO 2006). If total exposure is dominated by pollution in one microenvironment (e.g. ambient air) then controlling the pollutant in this microenvironment assures control of total exposure.

Exposure to air pollution at an individual level depends on many different factors, including the degree of activity that an individual undertakes, the time spent outdoors, and indoor sources. The

highest exposures to ambient air pollutants will be for those who are exercising or working vigorously outdoors and for children (because they have a higher breathing rate per kilogram than adults). In setting air quality standards, risk is assessed at a broad population level, and requires detailed information on the exposure of populations rather than individuals.

Information on people's time–activity patterns, based on detailed time–activity diaries, is often combined with information on air pollutant concentrations in the microenvironments where people spend their time to generate detailed exposure profiles. Time–activity patterns can be used to estimate total exposure, determine appropriate ways to reduce exposure and determine potential peak exposures in specific locations. Depending on the nature of the pollutant sources, a large proportion of daily exposure may occur in only a few hours, or exposures may be relatively well distributed throughout the day. In setting air quality standards, time–activity patterns can be used to determine the risk posed by specific microenvironments (e.g. ambient air), and can allow assessment of risk attributable to this exposure.

In assessing exposure for the setting of air quality standards it is necessary to determine the most appropriate averaging time for the health effects under consideration. For example, assessing exposure for benzene would involve calculating a long-term average as cancer is the most critical health outcome, whereas for sulfur dioxide the average may be in the order of minutes for broncho-constriction in asthmatics. For some pollutants multiple averaging periods may need to be considered as several health outcomes may be relevant.

Exposure to outdoor air pollution is most important for people or populations who spend a substantial amount of time outdoors. However, significant exposure to outdoor pollution can also occur due to outdoor air infiltrating indoors. For some pollutants (e.g. PM_{10}, carbon monoxide and $PM_{2.5}$), it has been estimated that about 90% of outdoor air infiltrates indoors. Therefore, exposure is dominated by outdoor air, even though people spend most of their time indoors. For reactive pollutants such as ozone and sulfur dioxide, exposure is dominated by time spent outdoors because these pollutants react rapidly with furnishings in the home, meaning that infiltration is low. Exposure to nitrogen dioxide can be significant from both indoor and outdoor sources, especially when gas is used for heating or cooking.

Ambient air quality standards are set on the basis of risk to the population rather than to individuals within the population and guide strategies for managing outdoor air quality. Thus, exposure assessments used for setting ambient air quality standards focus on exposure estimates for outdoor air quality, and consider exposure of populations only. Different levels of population exposure can be assessed, providing a range of risk estimates associated with each exposure scenario. This approach also enables the costs and benefits associated with the meeting of any new standards to be assessed.

An exposure assessment determines the frequency, extent and duration of exposure in the past, present and future by identifying air pollution levels, exposed populations, sensitive subgroups and potential exposure pathways.

4.2 OVERSEAS APPROACHES TO EXPOSURE ASSESSMENT FOR SETTING AIR QUALITY STANDARDS

Policy relating to ambient air pollution is developed and implemented on the basis of measurements made at fixed site monitors. Monitoring networks established in the United States and the European Union were established at locations chosen to measure variability in air pollution levels and thus enable population exposure to be determined. The WHO also recommends using monitoring to assess population exposure, and provides guidance on how that should be done. In contrast, air monitoring networks established in Australia for AAQ NEPM

monitoring have been established primarily at upper bound sites, with the aim of assessing compliance with air quality standards.

There are several approaches to assessing exposure and some include complex computer modelling. The approach used depends on the pollutant under consideration. Different computer models, of varying complexity, have been developed to estimate exposure; they can be used to estimate:

- local ambient concentrations, based on estimates of emissions from various point sources
- regional ambient concentrations, based on estimates of emissions, ambient monitoring data, the impacts of meteorology and interactions with other pollutants in the atmosphere
- likely individual exposures, using estimated ambient concentrations and an understanding of people's behavioural patterns that affect their exposure.

The section below provides an overview of international approaches to exposure assessment for use in setting air quality standards.

4.2.1 United States Environmental Protection Agency

4.2.1.1 Criteria pollutants

The US EPA uses slightly different approaches for each of the criteria pollutants, depending on the averaging period of concern with respect to health, significance of indoor sources in exposure and whether biomarkers (e.g. carbon monoxide) are used for the risk assessment. Only the inhalation pathway for exposure is considered, because it is most relevant to air quality standards.

The Air Pollution Exposure (APEX) model has been developed by the US EPA for determining exposure of the United States population to the criteria air pollutants. The model is available free of charge from the US EPA website.[4] The Hazardous Air Pollution Exposure Model (HAPEM), is a similar model that is used for air toxics (see assessment for air toxics, below). These models incorporate ambient monitoring data and time-activity data as the basis for exposure modelling, whereas in previous processes for setting standards for particles, ambient monitoring data have been used as the only measure of exposure. For other than criteria pollutants, ambient monitoring data are fed into the computer model with other data (e.g. time spent outdoors and infiltration rates to buildings) to assess total exposure and to identify exposure due to ambient air pollution. Scenarios are then modelled, taking into account potential changes to ambient air quality and resultant exposure profiles if potential air quality standards are met. The use of time–activity data allows exposure to ambient air to be separated from other sources, and thus allows assessment of the change in risk (e.g. number of deaths avoided) due to changes in ambient air quality to meet new air quality standards. The current review of the AAQ NEPM is giving consideration to incorporating time-activity data in an approach similar to that used for the other criteria pollutants in the United States (US EPA 2007).

For pollutants that have health effects associated with short-term averages (1-8 hours), time–activity data are important to account for short-term peaks in various environments (e.g. for nitrogen dioxide, short-term peaks due to cooking on gas stoves or spending time in motor vehicles).

For setting standards for nitrogen dioxide, sulfur dioxide and carbon monoxide, the US EPA uses the APEX model (and incorporates time–activity data) to estimate total exposure and take into account the contribution from various micro-environments (e.g. outdoors, homes, cars and work places). In setting standards for ozone, the US EPA uses a similar process, but focuses on time spent and activity outdoors, because exposure to ozone is almost entirely due to outdoor exposures.

[4] http://www.epa.gov/

For particles, the health effects that have been considered relevant to air quality standards are those for 24-hour or annual averages. Thus, ambient air quality data have been used, rather than detailed exposure modelling. For particles it is considered that ambient levels drive indoor exposures and, that ambient levels therefore adequately represent total exposure. In some cases this approach may underestimate or overestimate exposure; this is accounted for by listing this as a source of uncertainty in the subsequent risk assessment. More detailed exposure modelling for particles is now under consideration by the US EPA to improve estimates of exposure.

For carbon monoxide and lead, biomarkers are used as the basis of the air quality standards — carboxyhaemoglobin for carbon monoxide, and blood lead levels for lead. The exposure data for carbon monoxide is converted to a dose and then to a carboxyhaemoglobin level in blood. The health basis of the carbon monoxide standard is a carboxyhaemoglobin level in blood of less than 2 per cent. For lead, the 1998 standard is based on the NHMRC 1993 guidance (now rescinded) on blood lead levels of 10µg/dL. Complex pharmacokinetic modelling is used to convert ambient air quality data to exposure, then to dose and finally to blood lead levels.

In a risk assessment to determine health effects associated with any proposed standards, alternative scenarios are modelled in which exposure data are adjusted so that maximum values (or some statistical form, such as 98th percentile) meet the standard. The level of reduction is applied across the entire distribution of air quality exposure data. The assessment is conducted for pollution levels above those considered as background. Probabilistic models are used in some cases (e.g. for ozone).

Exposure is modelled and risk assessed on a selection of cities across the United States for which exposure–response functions for various health outcomes have been determined. In modelling of exposure for setting air quality standards, the US EPA considers the demographics of the population. Where particular groups of the population (e.g. children) are specifically considered for the basis of the standard, time–activity data for that particular group is incorporated into the models where appropriate.

For pollutants where detailed exposure modelling is conducted, the ambient air quality data are fed into the model, station by station. The exposure data are population weighted and specific to that population. For particles, the air quality data are averaged across the population. Exposure estimates are based on the exposure of the population not individuals. Time–activity data used is the profile for the population under consideration not individuals within the population.

Exposure data are generated to obtain an average representation of population exposure; therefore, peak-site monitoring data are excluded. Sites are selected to be representative of what the general population would be exposed to, and thus to be consistent with the air quality data used in the epidemiological studies from which the exposure–response functions have been determined. For ozone, all monitoring data are used because people's time-activity patterns are taken into account through detailed exposure modelling. However, inclusion of these data is considered not to have a major impact on the outcome of the exposure estimates.

In the United States, all air quality data used in exposure modelling for setting air quality standards come from fixed-site air monitoring stations. These stations have been established to measure population exposure. A full year's worth of data from each site is incorporated into the exposure model. With the APEX model, hourly air pollution data are required, and exposure for each hour of the year is calculated. Results are then averaged for the period appropriate for the health effects under consideration, and fed into the risk characterisation.

4.2.1.2 Air toxics

The US EPA has not set air quality standards for air toxics. The reference concentrations contained in the Integrated Risk Information System (IRIS) database are guideline values, and are based entirely on hazard assessment. A risk assessment has been conducted but not the risk characterisation stage (i.e. exposure and dose–response data have not been combined). Therefore, no exposure assessment is conducted as part of the RfC development.

The US EPA exposure model for air toxics (HAPEM) is used in evaluating the exposure as part of the evaluation of risk to the United States population from the 33 priority air toxics. The model predicts exposure using emissions inventory data and model outputs or ambient air quality data, where available. These data are then combined with hazard data to give a quantitative estimate of risk (e.g. increase in cancer).

Modelling data have been developed for each census tract in the United States. The data are generated for the centroid of each census tract and provide an estimate of what all people in the tract would generally be exposed to (as opposed to peak exposure). These data are input to the exposure model as a surrogate for ambient air quality data. Modelled data are compared with ambient monitoring (where available) to validate modelling outputs. Although not the case in the United States at present, the model could be used to estimate exposure for the purposes of setting air quality standards but would require the currently limited ambient air toxics data to validate the model.

Emissions inventory data from the toxic release inventory for all sources (mobile, industrial, area and mobile non-road) are used in the model. Meteorological and topographical data for each census tract are included, together with emissions data for that area. Similar emissions inventory data are available in Australia.

Like the APEX model for the criteria pollutants, the Assessment System for Population Exposure Nationwide (ASPEN) model has been developed to estimate total exposure to air toxics. Monitoring or modelled data are incorporated into the model, together with time–activity and micro-environmental data (e.g. type of heating at home) to give an estimate of total exposure to air toxics.

The data generated by both ASPEN and HAPEM are not consistent with the hazard data used in the risk assessment, because most hazard data for air toxics come from occupational settings or animal studies. However, because ambient air monitoring data for air toxics and epidemiological studies of populations in relation to ambient air are limited, the use of modelled data is considered appropriate for risk assessment. Sensitivity and uncertainty analyses are conducted and sensitive populations are considered in the risk assessment process. Only the inhalation pathway is assessed.

4.2.2 World Health Organization

The WHO establishes air quality guidelines for both criteria pollutants and air toxics based solely on health considerations. No quantitative risk assessment is conducted in developing the guidelines, but the WHO recommends that these considerations be taken into account when converting the guidelines to standards in individual countries.

The WHO provides guidance on how to establish monitoring to measure exposure of the population to the criteria pollutants. Issues to be considered in establishing monitoring are:

- where the population is located
- what the pollutant levels are and for how long populations are exposed
- what the areas and microenvironments are where the pollutant exposure is important.

Issues to be considered in determining the number of monitoring sites are:

- area to be covered
- spatial variability
- resources and instrumentation
- general population sites and hotspots.

A wide variety of monitoring locations are required to build a reasonable picture of exposure levels of the entire population. Therefore, the WHO considers it necessary to locate monitoring locations into the following areas:

- city or urban centre
- urban background
- suburban or residential
- kerbside or near road
- industrial
- rural
- other (e.g. microenvironments or location near sensitive subpopulations such as schools, hospitals and indoor environments).

The WHO also provides guidance on exposure assessment to be used in health impact assessments for policy development, public health decision making, research planning policy and establishment of environmental regulations (WHO 2001). The health impact assessment is required to estimate attributable cases, in an approach similar to that used by the US EPA. Coherence is needed between the measure of exposure used in the health impact assessment and that used to derive the exposure–response function and this is highlighted in the WHO guidance. The exposure assessment requires knowledge about the population exposure distribution, because the magnitude of the health hazard depends strongly on the level and range of exposure.

The WHO guidance is qualitative, and the organisation does not recommend specific modelling approaches. However, the WHO does highlight some key issues to be considered in the exposure assessment. The first issue is that the exposure-response functions are derived for a given exposure range and may only be valid for that experimental range. Any extrapolation beyond the experimental range needs to have the uncertainty of such an extrapolation clearly stated. In addition, at the lower end of the range a 'reference level' of exposure — below which no impact will be considered — may be part of the assumptions in the assessment. This reference level may refer to the natural or 'policy relevant' background levels of a pollutant such as ozone or particles. The following are typical situations where the 'reference level' may be of importance:

- exposure below some level may have no effect (i.e. threshold)
- the effect of exposure below some level may not yet have been assessed
- exposure below some level may be considered natural (i.e. natural background of ozone) and the respective impact may be excluded from the assessment
- exposure may be due to anthropogenic sources, but it may be impossible to reduce concentrations below a certain level because of background sources
- assessment of the impact of an environmental exposure above some defined level (e.g. air quality guideline) may be sufficient.

Another issue that the WHO identifies is the influence of the range of exposure experienced by the population under consideration. There may be circumstances where segments of the population are exposed to pollution levels higher than those observed in epidemiological studies that provide the exposure–response functions. The question is whether extrapolation of the exposure–response function to these levels is appropriate. It might be considered that the impact should only be quantified up to a certain level — for example, the maximum observed in the relevant

epidemiological studies — ignoring possible additional impact beyond these levels (WHO 2001). The WHO took this approach in developing the exposure–response functions for PM_{10} and $PM_{2.5}$ for the 2000 air quality guideline, where the range of pollution levels to which the functions applied was specified.

To calculate attributable risk, population-exposure distribution estimates are required. Exposure data can be obtained through a variety of methods, including:

- use of monitoring data
- use of monitoring data in spatially modelled air pollution levels (such as GIS maps); these data can be overlayed with population demographic data on population density to estimate exposure of the population
- a single estimate of average exposure; this may be warranted where air pollution levels are homogeneous across large areas and affect everyone.

The WHO (2000, 2006) strongly recommends that whatever exposure estimate is used, it must be the same as that used in the epidemiological studies from which the exposure–response functions have been obtained. For example, if an exposure–response function for PM_{10} has been obtained from studies using average ambient concentrations, then this measure should also be used in the health impact assessment. It is not appropriate to apply these exposure–response estimates to personal exposure measures.

In a recent review of the air quality guidelines, the WHO notes that the approach to exposure assessment depends on the purpose of the assessment (WHO 2006). Probabilistic modelling (such as that used by the US EPA) can be used to support policy-making and policy evaluation (such as the setting of air quality standards) by evaluating air pollution exposures in different (hypothetical) scenarios, population groups and locations. Data needs, costs and model complexity are less than are required by statistical empirical models (e.g. regression and factor analysis) or physical deterministic models (e.g. dispersion and mass balance models). The advantage of probabilistic models over these other models is that limitations of individual level data are overcome by using probability distributions of the input values. Input distributions incorporate the variability in the exposure data across as many individuals as possible. In addition to population averages, the predicted exposure distribution provides the range of exposures for the general population or subpopulation of interest, and the likelihood of exposures above a particular level (WHO 2006).

4.2.3 European Union

The European Union usually adopts the WHO air quality guidelines as limit values in their directives for air quality. Guidance is provided for establishing air monitoring networks for the purpose of estimating population exposure to air pollution. The number of monitoring stations is determined according to population density, using formulas that incorporate the number of people located in 'agglomeration areas'. For example, for PM_{10}, the European Union recommended the calculations shown in Table 4.1 to determine the number of monitoring stations required.

Table 4.1 Number of PM_{10} monitors required to estimate population exposure

Exposure scenario	Formula	Symbols
Urban background (per agglomeration)	$4\sqrt{I}$	I = population in millions
Background levels	A/50000	A = area of the country in square kilometres
Road sites	$3+\sqrt{P}$	P = population in millions

Industrial sites are measured such that there is sufficient coverage of industrial sources. The approach of measuring general and specific environments is used for all air pollutants for which standards, or limit values, have been set. In general, air monitoring stations are classified as urban, suburban or rural, with each being sub-classified again into traffic, industrial or background.

In the European Union, exposure modelling is used as part of the health impact assessment, in assessing the costs and benefits of actions taken to improve air quality. The integrated Regional Air Pollution Information and Simulation (RAINS) model is used, but no information is publicly available on the approach taken to the exposure assessment, data required and assumptions. The model has been developed to provide an integrated assessment of costs and benefits of air quality regulation in the European Union, including the health costs associated with exposure to ozone and particles. It is used not in developing limit values (standards) but in determining the costs and benefits through implementation of actions taken to meet the limit values.

4.2.4 United Kingdom

The terminology used in the United Kingdom for standards and objectives differs from that used in Australia and the United States. In the United Kingdom, a standard is a health guideline value (equivalent to the WHO guidelines) that is based solely on health considerations. Standards are converted to air quality objectives (equivalent to NEPM or United States standards) that take into account economic, social and environmental considerations, and have a number of exceedances and a timeframe for meeting the objectives. The objectives have a statutory role, and air management strategies must be developed to ensure that the objectives are met within the specified timeline.

The process of developing air quality standards and objectives in the United Kingdom does not include a formal quantitative risk assessment or risk characterisation. The Expert Panel on Air Quality Standards (EPAQS) conducts a hazard assessment and then uses professional judgement in deciding where the final standard is set. The Department of Environment, Food and Rural Affairs (DEFRA) evaluates the ability of the jurisdictions in the United Kingdom (i.e. England, Scotland, Northern Ireland and Wales) to meet the standards, and then sets the statutory objectives once the economic and social considerations have been considered. The United Kingdom approach to setting air quality standards and objectives is not transparent; it relies extensively on expert judgement, without providing documentation to support it. There is no formal exposure assessment as part of setting either air quality standards or objectives, and indoor, personal or total exposure is not taken into account. Ambient air quality data are used to assess whether a standard or objective is likely to be met; this is done by a simple comparison of the proposed standards against monitored levels.

Air quality modelling is used for mapping and for looking at different scenarios in the economic and feasibility analysis, but is not used for exposure or risk assessment in setting standards.

4.3 THE USE OF AIR MONITORING DATA FOR EXPOSURE ASSESSMENT IN STANDARD SETTING

As previously discussed, the exposure assessment process is the most critical part of any risk assessment. Approaches to estimating the exposure to air pollution range from averaging of the air quality data obtained from ambient air monitoring stations to complex analysis of time–activity patterns for the population, personal exposure assessments (including assessment of indoor air quality) and ambient air quality data, to obtain an accurate assessment of the exposure of a population to particular pollutants.

The approaches taken depend on the pollutant and how the health effects data have been obtained. For example in the United States, the exposure assessment for ozone involves an analysis of time–activity patterns, personal exposure, and estimates of both outdoor and indoor air quality. This complexity was thought necessary because the dose–response data for ozone were derived largely

from controlled human exposure and toxicological data, and exposure to ozone occurs almost entirely outdoors. In contrast, the exposure assessment for particles involved averaging ambient air quality data from 'average' sites, because the health data were derived from population based epidemiological studies, and infiltration of particles to indoor air is significant.

The exposure assessment for the 1998 AAQ NEPM used data from all ambient air monitoring stations and interpolated the upper percentiles for the pollutant levels, in an attempt to provide an estimate of the exposure across the population. The approach:

- did not provide an average estimate for exposure to the pollutants
- did not provide data compatible with the data used in the epidemiological studies to derive the exposure–response relationships
- was the same for each pollutant, which is not appropriate (RATF 2000).

When the NEPM was varied in 2003 to include an advisory reporting standard for $PM_{2.5}$, exposure assessment used in developing the standards was done by averaging the air pollution data across the monitoring networks. This is consistent with the approach used in the epidemiological studies that provided the exposure–response data for the risk assessment, and takes into account the significant infiltration of outdoor air indoors. The impact of alternative standards was assessed by a roll-back procedure whereby the air pollution distribution was reduced so that the peak values did not exceed a nominal standard. (See Appendix 2 for further explanation of this approach).

In any risk assessment the exposure assessment must mirror that used in the epidemiological studies used to derive the health endpoints and dose–response relationships. For ozone, this requires the use of time–activity data. Data from overseas for climates similar to Australia's (e.g. from California) could be used to fill data gaps in Australia. In 2003, the EPHC commissioned a national time–activity study to provide information to assist in exposure assessments for the review of the AAQ NEPM. The study focused on the most susceptible groups that are usually considered as a basis for air quality standards, and collected data from 4000 individuals in both summer and winter across all states and territories (see Section 4.4.3 for further information).

4.3.1 Australian ambient air quality data for exposure assessment

State and territory environment agencies have been monitoring air quality for many years, and there are extensive air monitoring networks in the larger cities. The review of the AAQ NEPM has found that the extent of monitoring varies in different airsheds, and that the monitoring sites put in place as part of the implementation of the NEPM are largely located to detect generally upper bound levels of pollution. In some jurisdictions, this type of monitoring dominates the data collected (NEPC 2007).

Implementation of the AAQ NEPM has led to consistency in monitoring methods used in Australia as well as an increase in monitoring in smaller jurisdictions and in regional centres. Monitoring of PM_{10} and $PM_{2.5}$ has increased nationally. The requirement for National Association of Testing Authorities (NATA) accreditation has ensured that good quality assurance and quality control procedures are in place.

The RAWG considered that if population exposure to air pollution was to be assessed from ambient monitoring data alone then monitoring stations needed to be located in areas of both high and low levels of pollution. The WHO, US EPA and the European Union all provide guidance on establishing air monitoring networks for this purpose. Because of their location, air monitoring networks in Australia do not generate data that allow assessment of the variability of the exposure of the population; the review of the AAQ NEPM is considering this issue.

The RAWG consulted widely with jurisdictions about whether the extent of monitoring being conducted for each of the criteria pollutants is sufficient to provide an estimate of population

exposure. In addition, the group reviewed the AAQ NEPM monitoring plans approved by the NEPC.

The group used the following criteria to assess the adequacy of the existing monitoring data for population exposure assessment:

- air monitoring should be conducted to comply with clause 14 of the AAQ NEPM regarding the number of performance monitoring stations for a region
- data should provide a cross section of exposures from high to low, to appropriately reflect the distribution of exposures across the community
- consideration should be given to distribution of sources for sulfur dioxide and lead, and to whether exposure arises mainly from point sources
- consideration should be given only to stations where at least 12 months continuous data are available.

The RAWG considered each pollutant separately, using the above criteria to assess the suitability of the data. On the basis of the information provided by jurisdictions and the review of the NEPM monitoring plans, the group concluded that exposure assessments and subsequent quantitative risk assessments would not be possible for smaller jurisdictions, because of the limited data available. Hence, in setting or reviewing national air quality standards under the NEPM, it would be necessary to extrapolate the results of any risk assessment process from larger jurisdictions to estimate the risk in smaller ones. With the implementation of the NEPM, this situation has changed for some pollutants — in particular, for particles. Most jurisdictions now monitor PM_{10}, and data are available to inform an exposure assessment process. The limitation is that the findings will reflect the upper bound of the pollutant distribution within the respective airsheds.

The amount of available data and the suitability of these data for use in an exposure assessment as part of standard setting process vary with the pollutant under consideration, as discussed below.

4.3.1.1 Sulfur dioxide and lead

Sources of sulfur dioxide and lead are now predominantly point sources, and current monitoring is designed to monitor the impact of those point sources on local air quality. The existing data are sufficient to characterise risk to any highly exposed groups near to those sources, but may not be sufficient to characterise the risks to the broader population. The data from point source monitoring are likely to lead to an overestimation of risks if extrapolated to the broader population. Since lead was phased out of petrol, levels of lead in ambient air have fallen below detectable levels in cities that are not influenced by industrial sources. Most jurisdictions have stopped monitoring lead, except near point sources.

The RAWG concluded that, to assess the risk to the population from exposure to sulfur dioxide, time–activity data would be required, because the short-term peaks are likely to be localised in nature rather than broadly distributed across the airshed. In addition, sulfur dioxide is a reactive gas that does not infiltrate significantly indoors, meaning that exposure to this pollutant occurs primarily from outdoor exposures.

4.3.1.2 Ozone and nitrogen dioxide

Although existing air monitoring networks monitor ozone at a significant number of sites in larger jurisdictions, monitoring is biased towards high-end exposures. Thus, if used alone, these data could lead to an overestimation of whole-of-population exposures and of subsequent risk. Overall, the RAWG considered that the current networks in larger jurisdictions were sufficient to provide data for exposure assessments, provided the bias of the data was taken into account in reporting the resultant risk estimates.

The relationship between ozone and nitrogen dioxide is complex. Nitrogen dioxide is a precursor to ozone formation; it also reacts with ozone, removing it from the air. Monitoring stations that record high ozone levels are therefore unlikely to record high nitrogen dioxide levels because the nitrogen dioxide is removed in the formation of ozone. Current monitoring networks often co-locate nitrogen dioxide monitors at peak ozone sites, to gain a better understanding of the chemistry of ozone formation within an airshed. Although this monitoring provides a good estimate of peak ozone levels, it may not provide an accurate representation of the distribution of nitrogen dioxide levels across an airshed and therefore of population exposures to the pollutant. This could lead to a tendency to underestimate nitrogen dioxide levels and the risk it poses. Nitrogen dioxide levels are highly variable across airsheds (US EPA 2008a). Epidemiological studies have found that levels of nitrogen dioxide measured in Australia are strongly associated with adverse health effects.

4.3.1.3 Carbon monoxide

Monitoring for carbon monoxide in limited in all jurisdictions, and the pollutant is not monitored at all in some smaller jurisdictions. The RAWG considered that the currently available data were generally insufficient to characterise population exposure. The group felt that this important issue should be addressed, given that there is increasing evidence from epidemiological studies (nationally and internationally) that significant health effects are observed at carbon monoxide levels below the current standard.

4.3.1.4 Particles

Monitoring of particles continues to provide significant challenges in most jurisdictions. Monitoring of PM_{10} is conducted primarily at upper bound sites; this may lead to overestimation of population exposure and risk, if used as the basis of an exposure assessment. Difficulties in monitoring of particles have led to some inconsistencies in data collected across jurisdictions. The loss of volatiles from the use of a tapered element oscillating microbalance (TEOM) for monitoring particles continues to be of concern, and may lead to underestimation of particle levels in some jurisdictions under certain meteorological conditions.

Monitoring of $PM_{2.5}$ is limited in most jurisdictions, and does not provide sufficient data to enable an accurate estimate of population exposure nationally. As part of the risk assessment for $PM_{2.5}$ for the AAQ NEPM, exposure was assessed by averaging the monitoring data across a city. This approach is used by the US EPA in developing air quality standards, and is consistent with the approach used in the epidemiological studies from which the exposure response data has been derived. In Australia, exposure was assessed in Brisbane, Melbourne, Perth and Sydney. Monitoring of $PM_{2.5}$ has decreased in some jurisdictions since that time.

4.3.2 Adequacy of Australian data for the setting of air quality standards

4.3.2.1 Criteria pollutants

Determining population exposure requires the establishment of monitoring stations to measure variability in air pollution levels (NEPC 2007). The review of the AAQ NEPM identified that the NEPM monitoring stations have been established for monitoring upper bound exposures not variability in pollution levels across an airshed (NEPC 2007). Clearly, within a region there will be a range of locations with high and low pollution levels to which individuals are exposed.

The RAWG concluded that any exposure assessment conducted as part of a standard setting process in Australia would rely on data from larger jurisdictions as there were insufficient data in the smaller jurisdictions. The group considered it generally acceptable to use data from the larger jurisdictions to determine the exposure of the Australian population, because more than two-thirds of the Australian population live in the major cities. However, there were some exceptions, based on knowledge of regional sources.

4.3.2.2 Sulfur dioxide

For sulfur dioxide, data from Brisbane, Melbourne and Sydney were considered likely to be representative of exposure of the Australian population residing in urban areas. However, the RAWG recommended the applicability of these assessments to Perth should be analysed further, because the sources located within the Perth airshed differ from those in other urban areas of Australia (RAWG 2002).

4.3.2.3 Nitrogen dioxide

The RAWG concluded that sufficient ambient monitoring data for nitrogen dioxide were available for Adelaide, Brisbane, Melbourne, Perth and Sydney to provide an assessment of population exposure for the purpose of setting standards. However, these data are often obtained at locations chosen to reflect peak ozone levels, and may not provide an accurate estimate of the distribution of nitrogen dioxide within an airshed. The RAWG recommended that the uncertainties in using such data be clearly identified and documented in the exposure and subsequent risk assessment, in setting air quality standards.

4.3.2.4 Ozone

For ozone, the RAWG concluded that there were sufficient ambient monitoring data available for Adelaide, Brisbane, Melbourne, Perth and Sydney to provide an adequate assessment of population exposure for the purpose of setting standards. However, these data may reflect high-end exposure and may lead to an overestimation of the risk posed. This uncertainty needs to be reflected in the outcomes of any risk assessment process.

4.3.2.5 Carbon monoxide

Data for carbon monoxide is limited in all jurisdictions. The RAWG concluded that there was sufficient ambient monitoring data available for Melbourne and Perth to provide an adequate assessment of population exposure for the purpose of setting standards. Given that motor vehicles are the main source of carbon monoxide in urban areas, the data from Melbourne and Perth should be sufficient for estimating exposure for the Australian population. Again, the fact that the air quality data are upper bound data needs to be acknowledged.

4.3.2.6 Particles

The RAWG concluded that sufficient ambient monitoring data were available for PM_{10} for Adelaide, Brisbane, Melbourne, Perth and Sydney to provide an adequate assessment of population exposure for setting standards, even though the data are upper bound. These five cities include most of the population of Australia; hence, the data should be sufficient to estimate population exposure. Continuing issues with particle monitoring — in particular, the adjustment of TEOM data for loss of volatiles — needs to be addressed and any data used must be consistent (e.g. all data need to be either adjusted or unadjusted). The exposure assessment needs to use consistent data to avoid bias.

Ambient monitoring data for $PM_{2.5}$ continues to be limited. At the time of the RAWG report, it was considered that data for Brisbane, Melbourne, Perth and Sydney were sufficient to provide a limited assessment of population exposure for setting standards. These data were used in the exposure assessment for developing the advisory reporting standards for $PM_{2.5}$. Since that time, $PM_{2.5}$ monitoring has been reduced in some jurisdictions, which may further limit any assessment of exposure of the population in some cities. However, there should still be sufficient data nationally to assess exposure at a national level. Data are now available from smaller jurisdictions.

4.3.2.7 Lead

Ambient lead levels are below detectable levels in most cities in Australia. It is therefore unlikely that an assessment of the exposure of the Australian population will be required for setting any

new standards in the AAQ NEPM. Where lead continues to be a point source issue, sufficient data are available to conduct an exposure assessment within the affected communities.

4.3.2.8 Air toxics

Air toxics come from a diverse range of sources including industrial processes, motor vehicles and domestic sources. Concentrations in ambient air are generally low, but highly variable across an airshed when compared with the criteria pollutants. In general elevated levels of air toxics will only be observed near sources. Given the variability in ambient concentrations of air toxics, certain groups within the population are more exposed and therefore at higher risk from air toxics than others. Exposure–response data (including unit risk factors) are derived largely from occupational or animal studies rather than population based epidemiological studies that are the primary source of exposure–response data for the criteria pollutants.

When the Air Toxics NEPM was made in 2004, it was acknowledged that data were too limited to conduct an exposure assessment for setting standards for these pollutants. Air toxics were not monitored routinely in Australia. The monitoring investigation levels in the Air Toxics NEPM were set solely on health considerations and are trigger levels for further investigation should these values be exceeded.

The goal of the Air Toxics NEPM is to collect sufficient data nationally to enable setting of air quality standards when the NEPM is reviewed in 2011. Data collection is focused on hotspots — for example, near roads, industrial complexes and in areas where wood smoke may be an issue. The monitoring is conducted at sites located to measure the cumulative impact of all sources in that area, not individual sources. Data are being collected under the Air Toxics NEPM; however, the available data are still limited and are currently insufficient to conduct an exposure assessment for setting air quality standards. Data being collected under the Air Toxics NEPM are 'peak' data and will only be applicable to communities in similar locations. Extrapolation across the entire population will lead to overestimation of risk. Modelling could be used to supplement the monitoring data as part of the exposure assessment.

4.4 SOURCES OF DATA FOR EXPOSURE ASSESSMENT IN AUSTRALIA

4.4.1 Ambient monitoring data

As discussed above, any exposure assessment in Australia will be limited by the availability of data, the bias toward high-end exposures and the lack of information on variability in air pollution levels experienced across the general population. This situation would need to be acknowledged in the uncertainty analysis of any exposure assessment. In addition, the exposure data will not match the data used in epidemiological studies when these use overseas exposure–response functions and population-based data. In addition, the assumption in the use of international exposure models is that the air pollution data inputs to the computer model reflect the variability in pollution experienced across the population. In some cases, peak data are excluded to obtain an average representation of what the population is exposed to. The impact of these issues on the uncertainly in the exposure modelling would need to be acknowledged in the risk estimates used in developing air quality standards.

4.4.2 Exposure models

Overseas computer models and approaches to exposure assessments can be used in Australia in setting air quality standards. NEPC used the US EPA approach to assess exposure for developing advisory reporting standards for $PM_{2.5}$. The approach mirrored that used to assess exposure in the epidemiological studies from which the health effects and exposure response data were derived.

Data for Brisbane, Melbourne, Perth and Sydney over a three-year period were available. These data had been obtained from TEOMs operated under similar conditions, with the same sample-

inlet heating temperature. These data were combined with population data obtained from the Australian Bureau of Statistics (ABS) — including information on sensitive populations, baseline health information and overseas exposure–response functions — to calculate the number of health outcomes avoided (e.g. avoided deaths) if air pollution levels were reduced so that peak levels did not exceed potential standards. This enabled cost and benefits to be calculated for each scenario. Air pollution data were averaged across the population and were not population weighted. The $PM_{2.5}$ approach highlighted the applicability and usefulness of adapting overseas approaches to exposure assessment for setting air quality standards in Australia.

Overseas models can be adapted for use in Australia where modelling of exposure is required, (as in the case of air toxics or ozone). However, the time–activity data, inventory, available air monitoring data and population demographics used in the model must be Australian. Models from areas with similar climatic conditions (e.g. California) should be used.

4.4.3 Time–activity data

In 2002 the EPHC funded the national Time–Activity Study (EPHC 2004) to inform the review of the AAQ NEPM standards. The study surveyed the activity of Australians in relation to their exposure to pollutants in the air, both outdoors and within their homes. A representative sample (approximately 4000 people) of children, young adults, and older Australians (aged 60 years and older) from all capital cities and two regional centres were surveyed. These age groups include those in whom respiratory and cardiovascular problems are most prevalent and are considered to be the most vulnerable to the effects of air pollution. This study surveyed these people to ascertain their time-activity patterns so that their level of exposure to environmental pollutants and the extent of respiratory problems with which this may be associated can be assessed.

The time–activity study surveyed activity of the same individuals in winter (September 2002) and summer (February 2003). Collecting data over two time periods allowed more information to be made available about annual cycles in activity and symptoms of cardiovascular and respiratory problems. The report highlights the key difference in findings between winter and summer.

Although limited in scope and number of participants, the study provides information on the time–activity patterns of the groups in the population who are more vulnerable to the effects of air pollution. These groups are likely to be the basis for the setting of air quality standards; thus, the time–activity data are an important input into any exposure assessment and will provide a mechanism to validate the use of overseas exposure models if required.

4.5 RECOMMENDED APPROACH FOR EXPOSURE ASSESSMENT

The recommendations for assessing exposure in setting air quality standards are:

- Where human epidemiological studies have been identified in the hazard assessment as key studies, air pollution data from fixed-site air monitoring networks should be used as the basis for the exposure assessment. Where population-based health data are used (e.g. information from time–series studies), air quality data for the whole airshed should be used, to be consistent with the health data.
- Air monitoring data for the criteria pollutants can be used as the basis for exposure assessment. Where data from smaller jurisdictions are not available, data from the larger Australian cities should be used to provide a national estimate for exposure. Uncertainties in extrapolating data from the larger jurisdictions need to be clearly documented.
- The exposure assessment should include scenario modelling, based on changes to the overall air pollution distribution, to allow determination of health outcomes that could be avoided if air pollution levels were lowered to meet any potential standard.

- Where fixed-site monitoring data are not available or are limited, air dispersion modelling should be used for the exposure assessment. For air toxics, the results of airshed modelling will be required, but any modelling data should be validated against ambient monitoring data. Air quality models are available for Australia and should be used in any modelling.
- Australia does not need to develop a computer model for exposure assessment. International models, such as those developed for California, can be used provided they make use of local data such as time–activity, meteorological data and air-pollution data.
- Where the hazard assessment has identified key studies that use individual rather than population-based health information, personal exposure data may be useful in determining the contribution of ambient air quality to overall exposure and risk. This information, combined with time–activity data, can be used (through modelling) in developing standards for ambient air quality or indoor air quality. The EPHC time–activity study provides information for the Australian population that can be used in exposure modelling for setting standards.

5 RISK CHARACTERISATION – PRINCIPLES AND GUIDANCE

This section provides firstly an overview of the essential requirements for a risk characterisation for air pollutants. This is followed by specific discussion of the process as applied to criteria pollutants (Sections 5.1 and 5.2) and the air toxics (Section 5.3). Detailed discussion of overseas approaches to risk characterisation for the criteria pollutants is provided in section 5.4. Uncertainty factors used as modifiers of the observed effect levels to protect susceptible sections of the population are described in section 5.5. The overall sources of uncertainty in the entire risk assessment process are discussed in section 5.6. Finally the recommended approach for risk characterisation is provided in section 5.7.

5.1 ESSENTIAL REQUIREMENTS FOR RISK CHARACTERISATION

Risk characterisation details the nature of potential effects on human health for the exposure conditions specified in the exposure assessment. It brings together knowledge about health impacts, exposure response and exposure, to calculate the likely number of people affected. The health effects range from mild (e.g. eye or throat irritation) to severe (e.g. increased deaths), depending on the concentration of the pollutant. The health endpoint to be considered must be identified in the hazard identification stage. The risk may be expressed, for example, as the annual number of attributable events (e.g. hospital admissions), or as the increase in cancer incidence or years of life lost.

5.1.1 Definitions of risk characterisation

The US EPA has published definitions and guidance for risk characterisation, which is defined as: 'A summary, integration, and evaluation of the major scientific evidence, reasoning and conclusions of a risk assessment. It is a concise description of the estimates of potential risk and the strengths and weaknesses of those estimates' (US EPA 1996).

Similarly, the European Union defines risk characterisation as: 'the estimation of the incidence and severity of the adverse effects likely to occur in a human population or environmental sphere due to actual or predicted exposure to a substance, and may include risk estimation, i.e. the quantification of that likelihood' (Hertel, 1996).

In Australia, enHealth provides guidance on risk characterisation; this guidance is consistent with approaches used in the United States and European Union (enHealth 2004). The risk characterisation should:

- identify the important strengths and uncertainties in the assessment, as part of a discussion of the confidence in the assessment
- include a full description of all elements of the assessment, drawing attention to the importance of the qualitative, as well as the quantitative, dimensions of the assessment
- include a summary of the key issues and conclusions of each of the other components of the risk assessment
- be transparent, with the conclusions drawn from the science identified separately from policy judgements, use of default values or methods, and clear articulation of assumptions used in the risk assessment.

5.1.2 Risk characterisation for air pollutants

Although risk characterisation is the last step in risk assessment, it is the starting point for risk management and the foundation for making regulatory decisions; for example, setting air quality standards. Risk characterisation:

- identifies and highlights the risk conclusions and related uncertainties of the risk assessment

- integrates the results from the hazard identification, exposure–response, and exposure assessments
- evaluates the overall quality of the assessment and the degree of confidence the authors have in the estimates of risk and conclusions drawn
- provides a qualitative or quantitative estimate, including uncertainties, of the nature, severity and potential incidence of effects in a given population, and summarises sources of uncertainty
- describes risks to individuals and populations in terms of extent and severity of probable harm
- communicates results of the risk assessment to the risk manager, by
  - allowing a synthesis of estimates of exposure levels and health risks
  - helping to identify the pollutant exposures that pose no significant health threat and those that present significant risks
  - indicating how risk varies with exposure, which is useful in evaluating a range of options for any potential air quality standards
  - helping to assess the costs and benefits associated with any potential air quality standards, to guide risk management decisions.

In setting air quality standards, risk characterisation applies to population rather than to individual risk. In theory, population risk can be calculated by summing the individual risks for all individuals within the subject population. Of course, this requires a great deal more information than is normally, if ever, available.

A probabilistic approach to determining population risk estimates the probable number of health-effect cases estimated in the population of interest over a specified time period. Population risk can be obtained by doing either of the following:

- summing the individual risks over all the individuals in the population; for example, using an estimated distribution of risk in the population, when such information is available
- using a risk model that assumes a linear non-threshold response to exposure, such as many carcinogenic models.

In these calculations, data will typically be available to address variability in individual exposures. If risk varies linearly with exposure, the number of cases can be estimated by multiplying the mean risk by the population size.

The results of risk assessment, as summarised in the risk characterisation, are but one consideration in the standard setting process.

5.1.3 Issues in risk characterisation

The following are the types of issues that need to be considered in risk characterisation for setting air quality standards:

- What is the risk (e.g. number of deaths or hospital admissions or increase in cancer risk) associated with any proposed standards?
- Does the risk assessment provide sufficient information to support the development of new air quality standards or changes to existing standards?
- What is the range of uncertainty around the estimated exposure level and the projected number of people who may be exposed to the pollutant?
- What data gaps are likely to affect the risk estimate or proposed air quality standards?
- Are studies being conducted that will provide new information that could fill one or more critical data gaps?
- Has the risk assessment been peer reviewed? If so, by whom, and what was the outcome of the review?

- Has a sensitivity analysis been conducted? If studies have been excluded from the hazard assessment, what is the impact on the results of the risk assessment? What was the rationale for excluding these studies?

5.1.4 Uncertainty and variability of the risk characterisation

Other questions primarily concern the issue of uncertainty (in assessing and characterising risk, the risk assessor must distinguish between variability and uncertainty). Variability in a risk characterisation arises from true heterogeneity in characteristics such as dose–response differences within a population, or differences in contaminant levels in the environment. The values of some variables used in an assessment change with time and space, or across the population whose exposure is being estimated. Assessments should address the resulting variability in exposure of the target population. Individual exposure, dose and risk can vary widely in a large population.

In contrast to variability, uncertainty represents lack of knowledge about factors such as adverse effects or contaminant levels; uncertainty may be reduced with additional study. Generally, risk assessments carry several categories of uncertainty, and each merits consideration.

5.1.5 Summarising the risk characterisation

Data lie on a continuum from strong evidence in humans (based on extensive epidemiology or other clinical or field observations) to weak evidence in humans, animals or other test systems (based on incomplete data in one or a limited number of species, or structure–activity relationships). Confidence in the conclusions of the risk assessment and the estimate of risk also lie on a continuum from high to low. The degree of confidence is largely based on the completeness, quality and consistency of the database (i.e. the weight of evidence). The following are issues that need to be considered through the risk assessment process that will ultimately affect any proposed standards.

- Where do the results of the risk assessment fit on the continuum from high to low confidence?
- Have exposures used in the risk assessment been measured in the population of interest? If exposures have been calculated through analogy, modelling or other estimation techniques, what evidence is there that the risk estimates are realistic?
- What is the degree of confidence in the existence of the risk and the magnitude of the risk estimate?
- If the risk is based on animal models, is there an observable parallel between humans and the positively responding animal species in terms of the absorption, metabolism, distribution and excretion of the pollutant of interest? If not, what is the basis for thinking such a parallel exists? Is there epidemiological evidence indicating that comparable effects seen in the animal model have been seen in human populations (e.g. heavily exposed occupational or environmental settings, and accidents)?
- Can population subgroups be identified who are at increased risk of exposure or especially sensitive to exposures, or both?
- At a given exposure or dose level, are there observable differences in the range of response among different human subgroups (e.g. infants, children, healthy adults and older adults)? If so, have these differences been evaluated and employed in the models used to calculate specific risks? If not, what evidence provides the basis for conclusions drawn about differences in sensitivity among subpopulations and potential risks?

Sections 3 and 4, above, outline the processes for assessing weight of evidence and identifying sensitive subgroups within the population that need to be considered when setting standards.

5.2 RISK CHARACTERISATION: CRITERIA POLLUTANTS BASED ON EPIDEMIOLOGY

As discussed earlier, the primary database for the health effects arising from exposure to the criteria pollutants is population-based epidemiological studies. The following, drawn primarily from the WHO (2000a), provides guidance on characterising risk using this type of data.

5.2.1 Specify the approach for obtaining measures of health outcomes

Estimating the impact of exposure requires information on the baseline occurrence (rate, prevalence) of one or more outcomes in the target population. Combined with the estimates of relative effect (relative risk or effect estimate), usually provided by epidemiological studies, information on baseline occurrence yields an estimate of impact of exposure in absolute terms; for example, the number of cases of disease or deaths. While exposure–response relationships may be derived from the international literature, the baseline disease occurrence is best obtained from data about the target population of the assessment. For example, if the standard setting process relied on epidemiological data for Melbourne and Sydney, then baseline health data should be obtained for those cities and used in the risk characterisation. If such data are unavailable, or inadequate, then health data from other populations can sometimes be used, with the potential limitations of such substitutions considered and thoroughly discussed.

5.2.2 Specify methods for estimating the number of attributable cases

The estimation of the burden of disease or mortality expected in the target population requires three basic elements, whose estimation is discussed above:

- distribution of the exposure in the target population
- estimates of the epidemiology-based exposure–effect function
- estimates of the baseline frequency of the health measure of interest.

Using this information, and assuming that exposure causes the health outcome, an epidemiology-based risk characterisation estimates the population attributable proportion (of disease or death) due to exposure. Population attributable risk is a measure that when applied to the target population, the population attributable proportion yields an estimate of the expected number of cases attributed to the exposure. (This measures is described more fully in standard epidemiological texts, for example, Rothman and Greenland 1998).

In practice, both the estimation and interpretation of the population attributable proportion, and its application to the target population, may involve a number of subtleties — for example, the choice of relative risk estimate or when there is evidence of confounding (see for example Greenland and Robins 1988, Rockhill et al 1998). The assumptions underlying the statistical methods used to estimate attributable proportions, or other measures of impact, and their implications for interpretation of those estimates, should be discussed.

The uncertainties in the data that contribute to the risk estimates, as well as any natural sources of heterogeneity in the effect of exposure, will often require the calculation of a range of estimates to fully describe the likely impact of exposure and to better reflect the uncertainty. This approach is often used by the US EPA in setting air quality standards, and was used in the development of the advisory reporting standards for $PM_{2.5}$ for the AAQ NEPM.

5.2.3 Issues in the interpretation of risk characterisation

The results of the risk characterisation require not only clear presentation, but also coherent interpretation, including explicit discussion of assumptions and limitations. Specific components of the overall uncertainty and their potential impact on the results ought to be addressed, as discussed above. Sensitivity analyses, in which the effects of key assumptions are explored quantitatively, may provide a better sense of the overall uncertainty of the estimates than purely qualitative discussion.

Risk characterisation aims to provide the number of cases or events attributable to air pollutants in a target population. When available, epidemiological studies play a central role in estimating attributable cases. In the absence of epidemiological data, epidemiological expertise is essential.

The step-by-step method described here for quantitative risk characterisation section is largely based on advice from the WHO guidelines (WHO 2000a):

Step 1 – Specify the purpose and framework of the risk characterisation.
Step 2 – Specify the method(s) used to quantify uncertainty.
Step 3 – Specify the measure(s) of exposure.
Step 4 – Specify the range of exposure to be considered.
Step 5 – Derive the population exposure distribution.
Step 6 – Specify the time window between exposure and effect.
Step 7 – Select appropriate health outcome(s).
Step 8 – Estimate the exposure–response relationship in the population of interest.
Step 9 – Derive population baseline frequency measures for the relevant health outcomes.
Step 10 – Calculate the number of attributable cases.

Steps 1 and 2 emphasise general conceptual considerations; Steps 3–5 relate to exposure; Step 6 also refers to health outcome; Steps 7–9 address health aspects; and Step 10 describes the final quantification process. The following section discusses Steps 3-10. Uncertainty is discussed in Sections 5.2.1 and 6.4 (Step 2).

5.2.3.1 Specify the measure of exposure (Step 3)

The measure of exposure must be defined. The risk characterisation requires information about the population exposure distribution, or, in the simplest case of dichotomous classification of exposure, the proportion of 'exposed'. Ideally, this exposure information would be available, but exposure may have to be estimated from available data, or obtained via an exposure assessment. If epidemiological evidence is used for risk characterisation, epidemiological exposure–response functions will be needed, to provide the quantitative association between exposure and health outcome. The definition of exposure in the two steps should be coherent. The population-exposure distribution for the exposure indicator or biomarker used in the epidemiological or experimental studies from which the exposure–response function was derived is needed for risk characterisation. There may be poor compatibility between exposure data used in epidemiological studies and the data on the population exposure distribution.

5.2.3.2 Specify the range of exposure to be considered (Step 4)

The magnitude of the impact of a pollutant on health depends strongly on the level and range of exposure. Risk characterisation uses this information to estimate attributable cases. Depending on the purpose of the risk characterisation, delimitations of considered exposure range may apply to either the lower or upper end of the exposure distribution, or to both. At the lower end of the range, a 'reference level' of exposure, below which no impact will be considered, may be part of the risk characterisation assumptions. The issue of 'reference levels' may have different aspects.

The following are typical situations where the 'reference level' may be of conceptual importance:

- Exposure below some level is assumed to have no measurable effect (threshold).
- Exposure below some level may be considered 'natural' (e.g. natural background level of tropospheric ozone) and the respective impact is excluded from the risk characterisation.
- Exposure may be due to anthropogenic pollution sources, but it may be impossible to attain concentrations below some level under the considered exposure reduction strategy.
- It may be necessary to assess only the impact of an environmental exposure above some defined level (e.g. above the air quality guideline level).

Depending on the perspective and the purpose of the risk characterisation, these aspects may be weighted in different ways. It should be clearly stated whether any reference level will be adopted, and why. The outcome of the risk characterisation is also influenced by the range of exposure observed in the population to which the risk characterisation is applied. It might be that a proportion of the population for which risk characterisation has to be conducted lives under exposure conditions that are higher than those observed in the epidemiological studies providing the exposure–response function. Thus, it is necessary to determine whether extrapolation of the exposure–response function up to these levels is adequate. It might be decided to quantify the impact up to a certain level — perhaps the maximum observed in the relevant epidemiological studies — ignoring possible additional impact beyond these levels. This is an issue that arises when exposure–response functions from overseas, or from areas with higher pollution levels, are used to characterise risk process for setting air quality standards.

The impact of assuming a threshold level is exemplified in the epidemiology-based impact assessment in Europe, from three countries, that estimated the number of cases of premature death, cardio-respiratory hospital admissions, incidence of chronic bronchitis in adults, acute bronchitis in children, restricted activity days and asthma attacks that could be attributed to PM_{10} exposure above 7.5 µg/m^3 (Künzli et al 1999, Seethaler 1999). A similar United States study did not quantify any impact for PM_{10} levels below 15 µg/m^3, and the study assumed a concentration of only 50 µg/m^3 for the population living in areas with PM_{10} annual mean levels higher than 50 µg/m^3. Obviously, these assumptions have a strong impact on the overall results.

5.2.3.3 Derive the population exposure distribution (Step 5)

To calculate the attributable number of cases, an estimation of the population exposure distribution is required. The availability of such information may be a limiting factor in risk characterisation. In some cases, the underlying epidemiological studies provide data about exposure distributions. In other situations, exposure data may be available from monitoring systems or other exposure-assessment studies. In sophisticated risk characterisation approaches, environmental monitoring data may be used to spatially model pollution levels. Such pollution maps may be combined with demographic data on population density to estimate the exposure (Filliger et al 1999). In other settings, a single estimate of the overall average exposure may suffice; for example, if concentration levels are fairly homogeneous across large areas and affect everybody (e.g. ambient outdoor air pollutants such as $PM_{2.5}$ or ozone). However, this approach may be more problematic for heterogeneous exposures.

The definition of exposure should be the same as that used in the epidemiological studies from which the exposure–response functions were derived. For example, if epidemiological studies present the risk function between ambient outdoor average concentrations of particulate matter and mortality (Katsouyanni et al 1997), risk characterisation requires the population exposure distribution for the ambient particle levels. It is not appropriate to apply these risk functions to the distribution of personal PM_{10} exposure (if available), which reflects both indoor and outdoor pollution with particles. Such personal exposure data would only be compatible with exposure–response functions from studies in which personal exposure and health effects have been assessed.

5.2.3.4 Specify the time window between exposure and effect (Step 6)

It should always be made clear whether the assessed health impact relates to immediate or delayed effects of exposure because the interpretations of the results (e.g. by policy-makers or economists) have to take the time window into account. When evidence suggests a time lag between exposure and outcome then studies that consider a lag effect should be given higher priority.

5.2.3.5 Select one or more appropriate health outcomes (Step 7)

As in the case of exposure, the health outcome measures have to be defined. Depending on the goal of the risk characterisation, the focus might be on only one or on several health effects,

ranging from minor health effects (e.g. respiratory irritation) through to death. If there is evidence for an environmental hazard being associated with several health effects, the impact should be assessed separately for each health endpoint. However, the selected endpoints may be overlapping measures, reflecting related aspects of a disease progression. For example, when morbidity and mortality are both considered, a single individual may be affected by both, first having disease (e.g. COPD) and then dying. As long as this possibility is acknowledged and described, risk characterisation may justifiably include evaluation of all considered endpoints. In effect, this is the explicit goal of 'burden-of-disease' approaches, such as disability-adjusted life years (DALYs), where time spent in ill health and premature mortality are combined in a composite index (Murray and Lopez 1996, de Hollander et al 1999).

In other cases of risk characterisation, in particular where an estimate of monetary cost is required, the assessment must clarify whether the health endpoints considered may be overlapping entities and whether this may cause 'double counting' of the same effects. For example, the tri-national air pollution impact assessment team in Europe (Seethaler 1999) decided to quantify the short-term effects of air pollution on hospital admissions, but not on emergency room visits. This was justified by the observation that, in many health-care systems, there is large overlap between the two measures of health-care system usage, because cardio-respiratory hospitalisations may often go through emergency rooms.

Considerations that will guide the selection of one or more health outcomes as the basis for risk characterisation include the purpose and use of the risk characterisation, the definition of 'exposure' and the availability of the respective data. For any of these aspects, assessors may decide not to include all measures of health in the risk characterisation for which epidemiological evidence is available. These decisions and the underlying arguments should be made explicit. Guidance on the selection of health endpoints for the setting of air quality standards is provided in Section 3.

5.2.3.6 Specify the exposure–response relationship (Step 8)

The quantitative association between the hazard and the outcome frequency is crucial for calculating the attributable number of cases. The exposure–response function is the key contribution of epidemiological studies to risk characterisation. It may be reported as the slope of a regression line, as a relative risk measure for a given change in exposure or as a comparison between 'exposed' and 'unexposed'. Due to the many sources of uncertainty in observational science, different epidemiological studies may lead to different exposure–response functions. Thus, for the risk characterisation, the process used to derive one or more exposure–response functions must be defined. The following issues have to be considered:

- Available epidemiological information should be systematically reviewed (as outlined in Section 3) to obtain information on reliable exposure–response relationships for every selected health outcome. The hazard identification process will normally provide an inventory of the relevant studies that are considered to be of acceptable quality. All studies that have quantitative information on exposure or that allow linkage to such information should be considered for the exposure–response evaluation. Sections 3.2 and 3.4 establish the process for evaluating epidemiological studies and determining exposure–response relationships.
- The process of combining studies for deriving an overall exposure–response relationship may be based on formal meta-analytic methods, pooled analyses or expert judgment (Blettner et al 1999). Published meta-analyses may also be useful, provided they are based on studies that are considered eligible for risk characterisation purposes. Measures of uncertainty around central point estimates should be derived, and information on heterogeneity between studies (e.g. from published meta-analyses) should be considered. Section 3 provides guidance on the use of meta-analysis.

- The studies selected during hazard identification may need to undergo an additional selection process, and may have to be weighted for the purpose of evaluating the exposure–response relationship for risk characterisation. In this case, the following should be taken into account:
  - The quality of exposure measurement needs to be considered.
  - Highest priority should be given to studies based on the same exposure metric as that used in the population for which the risk characterisation is required. Studies based on a different metric, but for which it is possible to convert results to the selected metric, should be given less weight.
  - Studies should also be evaluated on the basis of whether or not the estimated risks might apply to the population for which the risk characterisation is being conducted (i.e. generalisation from one to another population). For example, information on the possible presence of effect modifiers (e.g. local socioeconomic factors), or the importance of susceptible subgroups (e.g. asthmatics) that may drive the observed effects is valuable and should be taken into account.
  - The body of evidence may provide an estimated exposure–response relationship for a medium range of exposure levels, while risk characterisation may be required for a population mainly exposed to significantly lower or higher levels. Projecting exposure–response relationships beyond the range of exposure observed in the underlying studies normally involves uncertain extrapolations. The arguments for extrapolations, and their limitations or potential impacts need to be carefully addressed in the risk characterisation. Knowledge of the biological mechanisms underlying the specified effect may support the decision to extrapolate. In any case, allowance should be made for additional uncertainty.
  - The shape of the exposure–response function should be specifically evaluated in all available studies. Particularly, the possible existence of threshold levels ('no effect level') may be important for the risk characterisation.

5.2.3.7 Derive population baseline health measures for the health outcomes considered (Step 9)

In epidemiological studies, effects are most often reported as the relative change in risk rather than the absolute increase in number of subjects affected. Therefore, the risk characterisation process to quantify the impact requires information on the baseline occurrence (rate, prevalence) of the selected health outcome. With this information at hand, it is possible to calculate how many additional cases may be expected or may be attributed to some level of exposure.

While exposure–response relationships may be derived from the international literature, the baseline disease occurrence is best obtained from data about the population for which risk characterisation is being made. If such local data are not available, health-frequency data from other populations may sometimes be used. For example, if it is known that an environmental hazard increases the number of asthma attacks among asthmatics, quantification of the impact requires information about the number of asthmatics in the population, and about the average number of asthma attacks per asthma patient. Such information may be difficult to get for the target population, and data from other sources may be used (Künzli et al 1999). Apart from situations where complete population health measures are available (e.g. death statistics) baseline frequency data are estimates that are subject to errors and uncertainties.

5.2.3.8 Calculate the number of attributable cases (Step 10)

An epidemiology-based risk characterisation relies on the attributable risk concept or specific examples of impact assessment (e.g. Doll and Peto 1981, Rothman and Greenland 1998). This type of risk characterisation combines the major outputs from the steps described above, which are:

- estimates of the epidemiology-based exposure-effect function (i.e. mathematical link function between the degree of exposure and the expected change in health state)
- estimates of the epidemiology-based baseline frequency of the health measure of interest
- the distribution of the exposure in the target population.

5.2.3.9 Population statistics

Based on the observed frequency of health outcomes (incidence, prevalence) and the observed actual level of exposure, the expected number of cases will be calculated for an assumed baseline (or 'reference') level of exposure. This information needs to be combined with population statistics to provide an estimate of risk within a specified population. In theory, the (population) attributable proportion (AP) is the fraction of all cases attributed to a specified (dichotomous) exposure causing the health outcome:

$$AP = [p\ (RR - 1)] / [1 + p\ (RR - 1)]$$

where RR = relative risk for the health outcome due to the exposure, and p = the proportion exposed in the population.

The important assumptions at this stage are that there is a causal relationship between the exposure and the health outcome, that the relative risk estimate applies to all in the exposed group and that there is no confounding of the observed effect. The assumptions made and their expected impact on the impact estimates should be described in detail.

Consideration of uncertainties in risk estimates, and in exposure distribution, results in a range of impact estimates, rather than a single number. In more complex approaches, a probability distribution of impact is estimated. If one purpose of risk characterisation is to estimate the monetary cost of a public health impact, then economic valuation tools are needed. A sensitivity analysis may provide a better evaluation of the overall uncertainty of the risk estimates, and this should be done using the 95 per cent confidence interval associated with the exposure–response functions for the health outcomes under consideration.

5.3 RISK CHARACTERISATION: CRITERIA POLLUTANTS BASED ON CONTROLLED HUMAN EXPOSURE STUDIES

Results from controlled human exposure studies complement the findings of epidemiological studies. If epidemiological data are not available or are inadequate, then controlled human exposure studies can provide information on the mechanisms of action and the health effects of the pollutant in question.

The main parameters used to determine adverse health effects in controlled exposure studies include changes in lung function (e.g. decrements in forced expiratory volume [FEV1]), cellular and biochemical indicators of pulmonary inflammation (e.g. eosinophil and neutrophil concentrations in sputum), respiratory symptoms (e.g. increased cough, shortness of breath and chest pain on deep breath) and changes in bronchial responsiveness.

Many studies include healthy subjects in the 18–35 year age range, and include an exercise component to simulate likely exposure conditions for one of the potentially highest exposure population — that is, healthy workers carrying out high levels of activity. One limitation of these studies is that individuals are generally exposed to a constant level for the entire exposure period. This controlled-exposure scenario may be quite different to what might be expected in real-life situations, where a range of concentrations would generally be experienced over a specified time period. The response in the experimental setting is likely to differ from that experienced in real-world exposure conditions (US EPA 1996).

Another limitation of controlled-exposure studies is that the most sensitive individuals (e.g. people with severe asthma, or people with existing COPD or heart disease) are generally excluded from these experiments, limiting the generalisability of findings to the general population.

Controlled exposure studies are also usually limited to single pollutants. The results of these studies lead to the identification of an endpoint — NOAEL, LOAEL or, if exposure–response modelling is used, then a BMC (discussed further in Section 5.3) — for that specific study design and the group of subjects included in the study. The same NOAEL, LOAEL or BMC may not apply to sensitive groups. In many cases, these effect levels are based on a group mean response for healthy individuals; therefore, these studies may indicate large variability in individual response.

Because of these issues, the use of NOAEL, LOAEL or BMC in setting air quality standards requires the application of uncertainty factors, the choice of which will depend on factors such as:

- the size of the study population
- the strength of the study design
- the need to extrapolate from healthy individuals to sensitive populations
- the possible need to extrapolate from a LOAEL or BMC (these are both effect levels) to a NOAEL
- the quality of the database from which the NOAEL, LOAEL or BMC has been drawn (e.g. one study versus multiple studies).

The results of controlled exposure studies are only available for short-term health outcomes, usually for mild and reversible effects. For more severe or long-term effects, the results of epidemiological studies will be required.

A NOAEL may not be an estimate of a threshold level for adverse effects. It could be associated with a substantial (1–20%) but undetected incidence of adverse effects among the exposed population, or it could be much lower than a true population threshold (OEHHA 1999, US EPA 2004). This is because only a subset of individuals from the population has been observed, and because the experiment may not have been designed to observe all adverse effects associated with the pollutant. Therefore, it is not possible to safely conclude that the study concentration is not associated with any adverse effects.

5.4 Risk Characterisation for Air Toxics

As previously discussed, air toxics include a wide range of pollutants that are associated with health outcomes including:

- cancer
- reproductive and developmental effects
- respiratory effects, including those associated with asthma and chronic lung diseases, sensitisation and respiratory irritation
- central nervous system effects.

Internationally, ambient air quality standards are limited, although guidelines for control of emissions from industry or to assess risk in hotspots are available. The available information on the health effects of air toxics is derived from;

- occupational settings — these may include epidemiological studies of worker cohorts for large production volume chemicals
- toxicological studies of acute and chronic exposure to animals.

The HAPEM model — used to evaluate the risk to the United States population from the 33 priority air toxics — could be used to estimate exposure and could be combined with hazard data to set air quality standards, although this is not done at present.

5.4.1 Carcinogenic endpoints

Risk factors are generally used to assess carcinogenic risk. A unit risk factor is the risk of the incidence of cancer resulting from a lifetime exposure (usually 70 years) to 1 $\mu g/m^3$ of a carcinogenic substance. In deriving air quality standards from this approach, an 'acceptable' level

of risk is specified and the concentration of the pollutant corresponding to that level of risk is determined. This approach assumes a linear low-dose extrapolation; that is, no threshold for effect. This assumption is valid for genotoxic carcinogens (e.g. benzene) that react directly with DNA, but not for pollutants that cause cancer through other mechanisms (e.g. formaldehyde).

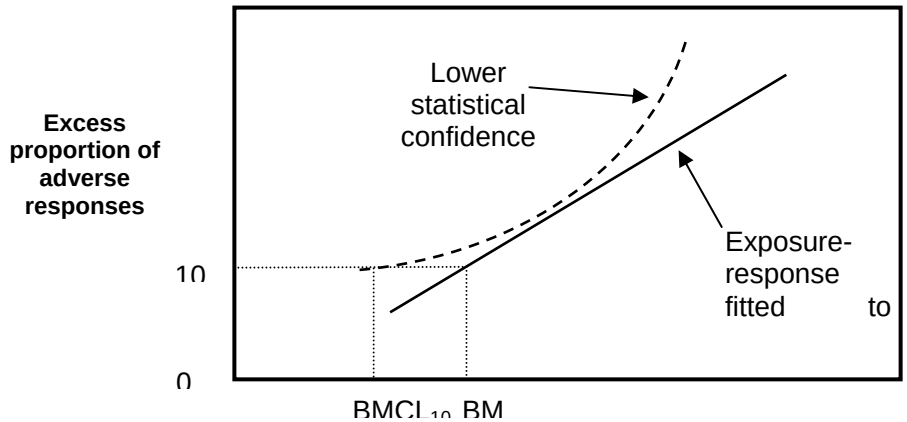
Bodies such as the WHO, US EPA, and OEHHA have developed unit risk factors for carcinogens; however the values can vary by an order of magnitude, depending on the studies used and the model chosen to derive the factor.

The benchmark-dose approach provides an alternative way of setting standards because it sets a threshold level associated with a low level of increased risk above background. This approach is used by US EPA (2000) in setting RfCs for non-carcinogenic pollutants and by Environment Canada in deriving air quality guidelines and standards. In the case of air pollutants, BMC is used in place of benchmark dose, because inhalation toxicology data are described in terms of air concentrations (OEHHA 1999, US EPA 2004). The term BMC is used in this document.

The BMC approach was developed for use primarily in toxicology, as a more quantitative alternative to exposure-response assessment than the NOAEL/LOAEL approach. While it could be used to develop standards for criteria pollutants, the substantial epidemiological databases generally available for these pollutants enable a much more sophisticated, as well as WHO-recommended, approach to standard setting. This approach estimates the population attributable proportion (of disease or death) due to the exposure i.e. an estimate of the expected number of cases attributed to the exposure.

The BMC is defined as 'the exposure that produces a predetermined change in response rate of an adverse effect compared to background'. It is estimated from a mathematical model fitted to the exposure-response data. To take the statistical uncertainties in the data into account, a confidence interval around the BMC is calculated. The lower 95% confidence limit of the BMC is often termed the BMCL. The BMCL may serve as a reference point or point of departure for deriving a health-based guidance level for human exposure. Figure 5.1, below, illustrates how a BMC (or benchmark dose) is derived. The BMC itself is sometimes used as the reference point rather than the BMCL; this approach has been called the *modified* benchmark dose or concentration. This distinguishes it from the US EPA methodology which uses the lower 95th percentile and confidence limits of dose rather than the central estimate. The reason for choosing the central estimates is because the lower bound is unduly influenced by the experimental protocols, including the sample size.

Figure 5.1 Derivation of benchmark concentration



The $BMCL_{10}$ in the case of figure 5.1 above illustrates the lower 95% confidence interval of the BMC_{10}, which is the concentration at which there is a 10% increase in adverse responses.

The BMC approach has been developed particularly by Crump (1984) and the US EPA (1995). It has been reviewed by Filipsson et al (2003), Travis et al (2005), and the United Kingdom Committee on Toxicity (COT 2007). The benchmark concentration is usually associated with a 1–10% increase in risk of an adverse effect occurring. The US EPA use a 10% response rate for BMCs when deriving chronic inhalation RfCs. OEHHA use a 5% response rate for calculating BMCs in the development of reference exposure levels, both acute and chronic (OEHHA 1999, 2000). Most examples in the literature use the BMCL when deriving health guidance levels.

Uncertainty or 'safety' factors to take account of types of studies used in toxicity assessments and the data available (e.g. occupational exposures, laboratory animal experiments, duration and levels of exposures) are applied to the BMC to derive a guideline or standard that would be protective of sensitive members of the general population.

Advantages of the BMC approach are that it:

- takes into account the entire exposure response curve rather than a single point as is done in the determination of the NOAEL
- accounts for the statistical power and quality of the data; wide confidence intervals lead to a lower BMC, reflecting the greater uncertainty of the database, whereas narrow confidence limits (reflecting better studies) result in higher BMCs
- is less influenced by the selection of doses as compared to a NOAEL, which is usually one of only a few preselected dose levels in many toxicity or population studies; the BMC can be calculated from datasets in which a NOAEL was not determined, eliminating the need for additional uncertainty factors to be applied to a LOAEL
- generally uses responses within or near the experimental dose range rather than relying on methods that require extrapolation to dose levels far below the experimental range. However, the BMC is not limited to one experimental exposure level, and the model can extrapolate outside of the experimental range.
- is consistent across a range of studies and health endpoints, for both carcinogens and non-carcinogens
- uses all relevant information.

In deriving a BMC, a mathematical model is selected, based on the data that are being analysed and the characteristics of the response. Simple and complex models are applied to limited and large databases, respectively. The model is applied with the appropriate considerations for statistical linkage, parameter estimation, and response. The dose–response model can be extrapolated below the biologically observable dose range enabling estimation of the response at specified (lower) dose levels as well as the dose corresponding to a specific response level (US EPA 2000, IPCS 2004).

If the experimental data available do not cover a sufficiently wide range of concentrations, it may not be possible to determine the shape of the dose-response curve and a BMC approach may not be feasible. In that case a linear approach using unit-risk factors should be applied in the derivation of the standard for air toxics.

In determining the approach to characterising the risk from carcinogenic substances, the mechanism for carcinogenesis should be taken into account. For genotoxic carcinogens that have a direct action on DNA, a linear approach using unit risk factors should be used. For pollutants that cause cancer through a multistage process, this needs to be factored into the standard setting process. For example, formaldehyde is known to cause nasopharyngeal cancer through a mechanism that involves initial irritation, which then leads to carcinogenic changes. In setting air quality standards, the mode of action of a pollutant needs to be considered.

5.4.2 Non-carcinogenic endpoints

Assessment of non-cancer health effects is usually based on the concept that a threshold concentration exists below which no adverse effects occur. While a threshold may be observed among individuals, the existence and magnitude of a population threshold, below which no members of the population would experience adverse effects, cannot be demonstrated (OEHHA 1999, 2000). Only a subset of the population is examined.

Much of the information on the health effects of air toxics is determined from animal or occupational exposure studies. Both the BMC and NOAEL approaches can be used to derive air quality standards and guidelines. As discussed in Section 5.5, uncertainty factors are applied to the LOAEL, NOAEL and BMC values from animal or human studies, to ensure that chronic and acute air quality standards are protective of human health for nearly all individuals.

5.4.3 Time conversion considerations

In developing standards for air toxics, the health outcomes associated with a range of exposure scenarios must be assessed. For example, the health effects associated with acute, short-term exposures to a pollutant may differ from those associated with chronic, long-term effects. The health outcomes assessed must be relevant to the exposure scenario and may lead to more than one standard for a pollutant. For example, in the air toxics NEPM, both toluene and xylenes have two standards — one related to acute exposures and their effect on the central nervous system and one accounting for longer term exposures resulting in reproductive and developmental effects. The basis of the standards should be the most sensitive health outcome for the exposure period under consideration. In determining the appropriate exposure periods, the approach used by the US EPA should be adopted. The definitions for the various durations of exposure are shown in Table 5.1.

Table 5.1 Definitions for various durations of exposure

Exposure classification	Exposure route	Duration of exposure
Acute	Oral, dermal or inhalation	Less than 24 hours
Short term	Repeated exposure by oral, dermal or inhalation route	More than 24 hours Less than 30 days
Longer term	Repeated exposure by oral, dermal or inhalation route	More than 30 days Less than 10% of lifespan in humans More than 30 days and up to 90 days in experimental animals
Chronic	Repeated exposure by oral, dermal or inhalation route	More than 10% of lifespan in humans More than 90 days and up to 2 years in experimental animals

Where available, inhalation studies should be used as the basis of air quality standards. Where inhalation data are not available, oral data can be used, but must be assessed for the relevancy of the health outcome for an inhalation exposure route. Additional uncertainty factors must be applied to account for the difference in route of exposure.

If standards are to be protective against chronic effects, continuous exposure must be taken into account. A conservative approach is used when setting air quality standards; the approach assumes that people may be exposed continuously for an entire lifetime, which is assumed to be 70 years. However, studies of adverse effects associated with long-term exposures of humans or animals generally involve discontinuous exposures. Common exposure scenarios are exposures for 6–8 hours per day, 5 days per week.

The default approach used by the US EPA and OEHHA to account for differences in effects associated with discontinuous and continuous inhalation exposures is an equivalent time-weighted average approach concentration (Cav), which uses the observed concentration (Cobs):

$$Cav = Cobs \times (\text{hours exposed}/24 \text{ hours}) \times (\text{days exposed}/7 \text{ days})$$

Where data has been derived from occupational settings, the following approach should be used to adjust from discontinuous to continuous exposures:

$$Cav = Cobs \times (10\ m^3/\text{day occupational exposure}/20 m^3/\text{day total exposure}) \times (\text{days exposed}/7 \text{ days})$$

Studies of adverse effects associated with exposures of humans or animals generally involve less-than-lifetime exposures, whereas air quality standards for chronic health outcomes are usually set to protect people over their entire lifetime. Typical experimental exposures periods are 5–20 years in an occupational setting, or exposures of experimental animals over about 10% of their lifetime (OEHHA 2000). The NEPC recommends that the approach adopted by the OEHHA be used to adjust from less-than lifetime exposures (sub chronic exposures):

- exposures of less than 8% of the expected lifetime use a 10-fold uncertainty factor
- exposures of 8–12 % of expected lifetime use a 3-fold uncertainty factor
- exposures greater than 12% of expected lifetime use a 1-fold uncertainty factor.

For a range of reasons, including practicability in monitoring a particular pollutant, a time conversion from the experimental time to another averaging period may be required. Guidance is provided by NHMRC (2006), OEHHA (1999, 2000) and US EPA on the application of Haber's Law for this purpose.

NHMRC recommend the following with respect to the application of Haber's Law:

- the starting point for setting standard averaging times should be the experimental data
- if standard averaging times are established on considerations other than deliberations of the biological and kinetic mechanisms giving rise to the health effect, it should be clearly stated that the averaging time does not reflect exposure times associated with health effects, and the reasons for this should be given
- adjustment of experimental exposure times to match a nominated averaging time can be achieved using the power version of Haber's Law
- timespan constraints are recommended, within which it would be considered appropriate to use Haber's Law
- for downward extrapolation from experimental exposures to a standard averaging time shorter than the experimental exposures, a value of n = 3 should be used for the exponential in the general form of Haber's Law; this is considered suitably precautionary for public health purposes
- for upward extrapolation from short experimental exposure times to a longer standard averaging time, a value of n = 1 should be used for the exponential.

5.5 OVERSEAS APPROACHES TO RISK CHARACTERISATION FOR CRITERIA AIR POLLUTANTS

The WHO, European Union and United Kingdom do not use quantitative risk assessment for deriving air quality guidelines, but do use it as part of cost–benefit analyses for exposure reduction processes and analysis of air management strategies. The WHO also provides guidance on how to conduct quantitative risk assessments, but not specifically for setting air quality standards.

The US EPA routinely undertakes quantitative risk characterisations, in particular in setting air quality standards for PM_{10}, $PM_{2.5}$ and ozone (e.g. US EPA 1996ab, 2004). The main approaches used are probabilistic. In 2001, the RATF recommended that a quantitative risk assessment should be used in setting air quality standards for $PM_{2.5}$, and recommended that the US EPA approach be used. The following sections outline the US EPA approach to risk characterisation, which was used in Australia to develop the $PM_{2.5}$ advisory reporting standards (see the case study in Appendix 2).

The US EPA quantitative risk assessments reports for each review of the standards are available on the US EPA website.[5] Key health endpoints used in these assessments include increases in daily and annual mortality (respiratory, cardiovascular and all cause), hospital admissions (asthma, COPD and cardiovascular disease) and decrements in lung function.

The risk characterisation is conducted for a limited number of cities where good air monitoring data and exposure–response functions are available. Cities are chosen to provide a wide range of exposures as well as source, pollutant mix and climate variability. For PM_{10} and $PM_{2.5}$, only ambient air quality data are used as part of the exposure assessment. For ozone, time–activity data are used with ambient air quality data in exposure models, because the predominant exposure to this pollutant is from time spent outdoors.

Scenario modelling is conducted to simulate the pollution distribution if pollution levels were reduced to meet a range of standards even at peak levels. An assessment of the risk as the number of attributable cases is produced for each scenario, based on the resultant exposure distribution. The risk characterisation provides a range of potential standards and the associated risk (as the number of cases; e.g. deaths or hospital admissions) that is considered by US EPA as part of the evidence for setting the air quality standards. The primary basis of health data is epidemiological, with toxicological data providing support for biological plausibility for the observed effect.

The risk assessment reports contain both a qualitative and, where possible, a quantitative assessment of uncertainty arising from all stages of the risk assessment process. The uncertainties considered are listed below:

- Accuracy of the estimates of the exposure–response functions (because these are empirically estimated there is uncertainty surrounding the estimates); if confounding variables are omitted, this could cause upward bias.
- Functional form of the exposure–response function — statistical significance of coefficients is an estimated exposure–response function; this does not necessarily mean that the mathematical form of the function is the best model of the exposure–response relationship.
- Transferability of exposure–response relationships — exposure–response functions may not be valid in times and places other than those in which they were estimated.
- Extrapolation of exposure–response relationships beyond observed data range — an exposure–response relationship estimated by an epidemiological study may not be valid at concentrations outside the range of concentrations observed in the study.
- Accuracy of air quality data (including consideration of composition and size for particles).
- Baseline health effects data — these may not be exactly appropriate for the epidemiological study or the risk assessment for various reasons (e.g. age-specific data may not be available in

[5] http://www.epa.gov/ttn/NAAQS

some locations). Baseline incidence may change over time for reasons unrelated to air pollution.

- Sensitive subgroups — populations in locations used for the risk characterisation may have more or fewer members of sensitive subgroups than the locations where the exposure–response functions have been derived. Thus, the functions may not be appropriate.
- Omitted effects — some health effects attributable to air pollution may have been omitted.

Susceptible populations considered included children and older adults. In the United States, legislation requires special consideration of children in assessing and monitoring air quality and in the development of air quality standards. In California, air monitoring networks must be set up to measure exposure of infants and children to air pollution. Also, in both California and the United States as a whole, all air quality standards must be assessed to ensure protection of children's health. The legislation identifies the following issues as needing to be addressed when assessing the health impacts of air pollution:

- children have narrower airways than adults; thus, irritation or inflammation by environmental factors such as air pollution may be mild in adults but could result in a potentially significant obstruction of the airway in a young child
- children's ventilation rates and the surface area of their lungs differ from adults and make them more susceptible to the effects of air pollution
- developing organs and tissues of infants and children are more susceptible to damage from some environmental contaminants than are adult organs and tissues
- exposure patterns for children may be different to those of adults, leading to disproportionately high levels of exposure in comparison to the general population; children spend significantly more time outdoors than adults
- air pollution exacerbates asthma, which is particularly prevalent in children.

5.6 UNCERTAINTY FACTORS

Uncertainty factors are numerical corrections applied to effect levels derived from limited experimental or epidemiological data (e.g. NOAEL, LOAEL and BMC), to derive air quality standards or guidelines for a general population that includes particularly vulnerable members such as the elderly, children or people with pre-existing disease.

In the absence of comprehensive population-based exposure–response data, mathematical modelling of exposure–response relationships for air pollutants can be used for standard setting (e.g. to derive a BMC). The modelling is highly dependent upon the quality and extent of available toxicological or epidemiological data, but often these are inadequate for accurate modelling because the studies were designed for hazard identification rather than for determination of exposure–response. Where exposure–response modelling is not possible, air quality standards and guidelines are often derived from the NOAEL or LOAEL. The advantage of using a BMC over a NOAEL approach is that the former takes into account the dose–response data, whereas use of a NOAEL does not explicitly incorporate this data. Where sufficient data are available to determine BMCs, this approach should be used in preference to the use of a NOAEL; however, such data are limited and will not be available for all pollutants.

Historically, chemical risk assessment has used a value of 100 as the default uncertainty factor to derive a reference dose for the general population, based on a NOAEL or LOAEL from a chronic study in animals. For the criteria pollutants, the choice of uncertainty factors for deriving air quality standards from a NOAEL or LOAEL is not as clearly defined as it is for deriving guidance values for chemicals. The default 100-fold value represents the product of two factors of 10, which allow for interspecies differences and human variability. Uncertainty can be reduced by more data. For example, the use of controlled human exposure studies eliminates the 10-fold interspecies extrapolation uncertainty factor.

Uncertainty can also be reduced by the use of chemical specific adjustment factors (CSAFs), provided appropriate data are available. The IPCS has produced a guidance document on the application of the default 100-fold safety factor. The guidance details the development of CSAFs for interspecies and human variability (IPCS 2005) allowing risk assessors to move away from the default if adequate data are available. CSAFs provide a way to incorporate biologically based quantitative data on interspecies differences or human variability.

In the absence of comprehensive data, additional uncertainty factors may be applied if assessors consider that in a particular risk assessment additional uncertainties exist. Some of these have been discussed in Section 5.2; they include:

- the size of the study population
- the strength of the study design
- the need to extrapolate from healthy individuals to sensitive populations
- the need to extrapolate from a LOAEL or BMC (these are both effect levels) to a no-effect level
- the quality of the database from which the NOAEL, LOAEL or BMC, or the health effects data have been drawn (e.g. one study versus multiple studies).

Uncertainty factors are sometimes used to allow for cross-route uncertainties (e.g. an oral exposure study can be used for air quality criteria), database deficiencies and the severity and irreversibility of effects. The choice of uncertainty factors is based on expert judgement.

The total uncertainty factor, taking into account all these considerations, can be up to 100 000. If such uncertainty exists in the evidence for a health effect, then it should be considered that there is insufficient evidence to support the development of an air quality standard for that pollutant.

Uncertainty can be reduced when risk assessors have new data to allow them to develop a quantitative risk assessment and reduce the uncertainty incorporated into previous risk assessments. In the case of a pollutant where the results of the epidemiological studies indicate that effects are observed in the community below the NOAEL or LOAEL obtained from controlled exposure studies, then an additional uncertainty factor would normally be applied. This case will arise only when the epidemiological studies are of insufficient quality to provide reliable exposure–response functions.

As an example, when the AAQ NEPM was made in 1998, the standard for nitrogen dioxide was based on a NOAEL from a controlled exposure that was identified for healthy people. A safety factor was applied to the NOAEL to obtain the standard. At that time, the results of epidemiological studies were conflicting, and the exposure–response data were too unreliable to derive a standard through a quantitative risk assessment process. Since that time, a large number of studies have been conducted worldwide, from which exposure–response functions have been derived. This includes studies in Australia that have found strong and consistent associations between exposure to nitrogen dioxide and adverse health outcomes. For the review of the standards in the AAQ NEPM or the development of new standards, reliable exposure–response functions are available and should be used in the risk assessment process outlined in this document.

5.6.1 OEHHA guidance on uncertainty factors

The OEHHA applies the uncertainty factors shown in Table 5.2 to the derivation of reference exposure levels for (threshold) air pollutants (OEHHA 1999). This approach is consistent with that used by the US EPA in the derivation of reference concentrations for air toxics.

Table 5.2 Uncertainty factors in deriving reference exposure levels for (threshold) air pollutants (OEHHA 1999)

Methodology and study type	Nature of uncertainty	Uncertainty factor
Benchmark concentration, animal study	Interspecies Intraspecies	3 10
Benchmark concentration, human study	Interspecies Intraspecies — study includes sensitive subjects Intraspecies — study does not include sensitive subjects	1 1 3 or 10
NOAEL, animal study	Interspecies Intraspecies	3 or 10 10
NOAEL, human study	Interspecies Intraspecies — study includes sensitive subjects Intraspecies — study does not include sensitive subjects	1 1 10
NOAEL, any study	LOAEL — mild effects LOAEL — other than mild effects	6 10

A different scheme is used for the criteria pollutants, based on the use of epidemiological data, as described in Section 5.1. Additional uncertainty might result from the size of the study population or the strength of the study design. Other factors discussed in Section 5.2, such as the quality of the database, are not considered by the OEHHA methodology.

5.6.2 NHMRC guidance on uncertainty factors

The NHMRC has provided guidance on the use of the uncertainty factors in setting air quality standards (NHMRC 2006):

- The possibility of sensitive sub-populations within the general population should be considered when setting an air quality standard. However, the default composite uncertainty factor of 10 for human variability is considered to account for sensitive persons within the general population exposure–response distribution.
- Where possible, CSAFs should be developed. Where this is not possible, then the default composite uncertainty factor of 10 for human variability should be adopted.
- In the absence of information showing that children have increased sensitivity to a particular air pollutant, no adjustment for child exposure is needed because the default adjustments for human variability within the adult population adequately protect children as well.

5.7 CHARACTERISING SOURCES OF UNCERTAINTY

Particularly critical to full characterisation of risk is discussion of the uncertainty in the overall assessment and in each of its components. The uncertainty discussion is important because:

- information from different sources carries different kinds of uncertainty, and knowledge of these differences is important when uncertainties are combined for characterising risk
- a discussion of the uncertainties will help to identify where additional information could contribute significantly to reducing uncertainties in 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.

A discussion of uncertainty requires comment on issues such as the quality and quantity of available data, gaps in the data base for specific pollutants, quality of the measured data, use of default assumptions, incomplete understanding of general biological phenomena and scientific judgments or science policy positions that were employed to bridge information gaps.

Measurement uncertainty refers to the usual error that accompanies scientific measurements–standard statistical techniques can often be used to express measurement uncertainty. A substantial amount of uncertainty is often inherent in environmental sampling, and assessments should address these uncertainties. Similarly, there are uncertainties associated with the use of scientific models (e.g. dose–response models, and models of environmental fate and transport). In evaluating model uncertainty, the scientific basis for the model and available empirical validation should be considered.

Uncertainty also stems from data gaps — both hazard and exposure data — that require estimates or assumptions used in the assessment. The data gap is often broad; for example, the absence of information on the effects of exposure to a pollutant on humans or on the biological mechanism of action of a pollutant. A statement of confidence should be included that reflects the degree to which the risk assessor believes the estimates or assumptions adequately fill the data gap.

Qualitative information on methodology, alternative interpretations and working assumptions (including defaults) is an important component of risk characterisation. For example, specifying that animal rather than human studies were used in an assessment makes clear that the risk estimate is based on assumptions about human response to a particular pollutant rather than human data.

5.8 RECOMMENDED APPROACH FOR RISK CHARACTERISATION

The recommendations for characterising risk as part of the process of setting air quality standards in Australia are:

- The probabilistic approach used by the US EPA and by the NEPC in developing standards for $PM_{2.5}$ — in which a range of scenarios are assessed for a range of health outcomes — should be used. This approach transparently provides information on the risk associated with potential standards and makes it possible to evaluate the costs and benefits of alternative scenarios.
- The approach used by the WHO to guide the use and interpretation of epidemiological studies for risk characterisation for criteria pollutants should be adopted.
- Where toxicological data are used as the basis for the hazard assessment, inhalation studies should be used where possible. Where such studies are not available, oral studies can be used but must be assessed for the relevancy of the health outcome for inhalation exposure. Additional safety factors must be used to account for differing routes of exposure.
- For air toxics, both carcinogenic and non-carcinogenic, a BMC approach should be used, provided that data are available to support the application of this approach. In setting air quality standards for carcinogens, the mode of action for the causation of cancer must be taken into account and will determine the approach to setting the standard. For genotoxic substances, a linear model using unit risk factors should be used.
- The NHMRC (2006) guidance on the choice and application of uncertainty factors should be adopted in the risk characterisation for non-criteria pollutants. The OEHHA approach may provide further guidance on the use of uncertainty factors.
- Health outcomes for air toxics may vary for a range of exposure outcomes, and a number of standards may be required for a given pollutant. The most sensitive health outcome for the relevant exposure period should be used as the basis for the standard. In defining exposure periods, the approach used by the US EPA should be adopted.
- Conversion of averaging times from experimental exposure times to averaging periods appropriate for air quality standards should be done using Haber's Law, following the guidance set out in Section 5.3.3.

6 RANGE OF POTENTIAL RISK BASED GUIDELINES

This section summarises the pertinent points from the hazard assessment section that feed directly into risk characterisation. It then proceeds to set out the central importance of the risk characterisation for the range of potential risk based guidelines (box 3 of the framework).

As set out in Figure 2.3, the risk assessment process as outlined in sections 3-5 of this document are a critical input to the setting of air quality standards. However, other considerations must be taken into account in developing standards in the NEPC context. These include science and policy, including social issues as set out in Section 6, as well as economic considerations. All inputs are considered before a decision is made by NEPC on the final standards.

The risk assessment process leads to a range of potential risk-based guidelines that are then further assessed through the consideration of environmental, social and economic factors that are addressed through the Impact Statement that accompanies any draft NEPM or NEPM standards.

The risk assessment must clearly identify the health outcomes and sensitive groups that were considered as the basis for any air quality standards as well as the relevant exposure response functions that were used in the risk characterisation. The averaging periods to which these data apply also need to be clearly identified. Only health outcomes and sensitive groups that have been identified, through the weight of evidence analysis, to be associated with evidence that an observed association with a pollutant is either sufficient to infer a causal relationship or sufficient to infer a likely causal relationship (i.e. more likely than not) should be considered in the risk characterisation process.

Appropriate exposure data, whether obtained from ambient monitoring data or through modelling, needs to be identified and used in the exposure assessment. All assumptions made in the estimates of exposure must be clearly identified. A range of exposures need to be considered including current air quality and what could be achieved through air quality management strategies. For pollutants that have a naturally occurring background levels, such as ozone or particles, this needs to be taken into account when identifying the lower range of the exposure scenarios to be considered.

In identifying the health outcomes and sensitive groups, issues such as the precautionary principle and environmental equity need to be considered to ensure that the most vulnerable groups in Australia, where possible, are assessed in the development of the standards. If data from Australia are not available but there is a significant body of work overseas that indicates that there are critical health outcomes or sensitive groups that might be impacted by a particular pollutant then application of the precautionary principle, as required under the NEPC Act, would require that the international data be used as part of the standard setting process.

The risk characterisation should be conducted for each of the health outcomes and sensitive groups identified using the relevant exposure-response functions and exposure data. The outcomes of the risk characterisation can be expressed as either attributable risk or as the number of health outcomes avoided, as was the case in the development of $PM_{2.5}$ standards as set out in Section 5.7. In developing the standards for $PM_{2.5}$ the steps outlined in Section 5 were followed and should be utilised in future setting of air quality standards. A probabilistic approach should be used to assess the health impact of different air quality scenarios that would be required to achieve potential risk based standards for each health outcome and sensitive groups that have been identified through the hazard assessment.

The results of the risk characterisation provide a range of estimates for each of the health outcomes and sensitive groups associated with a particular exposure scenario and the health benefit

(reduction in attributable cases or number of health outcomes avoided) linked to a reduction in the level of the pollutant under consideration. This information is the basis of the range of potential risk based guidelines referred to in Box 3 of Figure 2.3. This data should be tabulated and can be used in a cost-benefit analysis to inform the decision on the final standards adopted by NEPC.

In deciding on where the final standards should be set NEPC considers a range of factors including:

1. The health benefit associated with reducing air pollution levels to meet potential standards (avoided health outcomes and associated costs)
2. The range of management strategies that would be required to reduce air pollution levels to meet the standards and the costs associated with those strategies
3. The costs of monitoring for the pollutants to which the standards would apply
4. Any social or regional differences, including cultural practises, that might be impacted by strategies implemented to meet the standards

The outcomes of the process outlined in this document are critical input to the final decision but are not the only considerations. All factors must be weighed up and the overall balance to adopting new standards must be a benefit. Input from public consultation is also an important input to the standard setting process to ensure that all stakeholders' views are considered in weighing the information to derive a final standard.

7 POLICY CONSIDERATIONS

7.1 CARCINOGENS — 'ACCEPTABLE' LEVEL OF RISK

The issue of acceptable risk has often been discussed in the context of environment and health standard setting in Australia. However, an agreed numerical value has not previously been expressed for a level of 'acceptable cancer risk', as a firm guideline value.

Based on feedback arising from previous community consultations as part of NEPM processes, a 'ceiling limit' is that cancer risk at the population level should not exceed 1 in 100 000 (1×10^{-5}). This appears to be the maximum acceptable risk for the Australian population. Thus, for setting air quality standards, the estimated risk from population exposures should not exceed 1 additional case per 100 000 of the population per year. An exposure-reduction approach presupposes that, where possible, efforts should be made to reduce the risk to lower levels.

For most of the criteria pollutants, epidemiological studies have been unable to identify a threshold below which adverse health effects do not occur. This is also the case for many air toxics, especially carcinogenic substances. The lack of an observed threshold for effect means that any air quality standard will have some level of associated risk, and that standards can never be fully protective of health — they will be risk-based standards.

In setting air quality standards, the level of health protection for a given level is balanced against the costs and benefits associated with meeting that standard. The final standard is set at the level that encompasses the best overall outcome of this analysis. Achievability of meeting a standard must also be taken into account. For example, if background pollutant levels (e.g. of ozone and PM_{10}) from natural sources are high, then in setting a standard for these pollutants, the assessment of risk and associated costs and benefits must include the anthropogenic component. The aim in setting the standard is to minimise the risk while balancing the costs, benefits and achievability of meeting the standard. These issues should be clearly documented in the impact statement that accompanies draft standards, to ensure transparency in the decision-making process.

7.2 IDENTIFYING FACTORS FOR COST-BENEFIT ANALYSIS

The development of standards per se will not specifically result in any costs or benefits, but the implementation of management strategies to achieve the standard will. Under the NEPC Act, the costs and benefits of any proposed management strategy must be assessed to guide the decision on making or varying a NEPM. In developing or revising air quality standards, the costs and benefits of meeting a range of standards should be considered. This analysis is undertaken as part of the impact statement. An impact statement addresses the financial, health and environment impacts of the proposed actions or program, demonstrates that the proposal is justified, and provides a reasonable basis for informed comment by stakeholders and the community.

Before a cost–benefit analysis can be conducted, the factors to be considered need to be identified and quantified. It is important to consider and quantify the direct costs of implementing management strategies and air quality monitoring. Most of these actions are likely to fall to government environmental agencies, which will be best placed to calculate the implementation costs on a pollutant-by-pollutant basis. Where there is the potential for indirect costs to flow on to industry and the broader community, these should be estimated in consultation with the affected stakeholders. The benefits of achieving air quality standards are normally quantified by predicting or measuring key human health outcomes, which can be converted to 'health costs avoided' where cost estimates are available. This analysis is complicated by the fact that the sources of health statistical data are not always clear, and broad consensus about the cost of key health outcomes is often lacking.

This section attempts to resolve some of the issues mentioned above by discussing where health statistics and costs can be obtained to feed into a cost–benefit analysis. It does not provide a method for analysing costs and benefits, but simply provides guidance on where to find information to inform the estimates of the health costs.

7.2.1 Obtaining health statistics

Key health outcomes that need to be considered as the basis for setting air quality standards include:

- mortality
- hospital admissions
- emergency department visits
- exacerbation of asthma and other respiratory or cardiovascular diseases
- decrements in lung function.

Where data are available, broader quality of life issues (e.g. work-loss days and school absenteeism) should be considered.

Mortality, hospital admissions and emergency attendance data for various ICD codes can be obtained from a variety of sources including the ABS, the AIHW and jurisdictional health departments. Data are readily available at a broad population level for different age groups, sex and ICD codes. For setting air quality standards, broad population-based data for different age groups, sex and ICD codes are sufficient, because they are more applicable to the whole population. This information will enable assessment of vulnerable groups as the basis for air quality standards if exposure–response data are available.

Records of general practitioner (GP) attendances for specific ICD codes are not easily obtainable. Some areas have GP sentinel networks, but these only register a limited number of GPs. At present, it is not possible to obtain the baseline health statistics needed to assess and cost this outcome. A watching brief should be maintained on whether a state or national surveillance system for GP attendances is developed. If such data do become available, then this aspect should be assessed as a critical outcome in setting air quality standards.

7.2.2 Obtaining health costs

The association between air pollution and increases in mortality (both short term and long term) is a key driver for setting air quality standards. Assessing the costs associated with loss of life has often been the subject of significant debate, and has led international agencies to look at alternative ways of assessing the costs associated with increased mortality. The two key approaches commonly used are:

- estimating the number of deaths associated with alternative standards
- estimating numbers of years of life lost.

In May 2007, the Office of the Australian Safety and Compensation Council (ASCC) commissioned Access Economics to review international approaches to valuing life. The aim was to derive low, base and high values for the value of a statistical life (VSL) and the value of a statistical life lost (VSLL) (Access Economics 2008), for use as inputs in analysing cost effectiveness cost benefit. VSL is the marginal dollar value of a human life; a VSLY is the marginal dollar value of a year of healthy human life.

A meta-analysis of published figures for VSL from around the world yielded an average value for VSL of $6 million in 2006 Australian dollars, with a range of $5 million to $7.1 million (Access Economics 2008). These estimates are not age specific. Due to significant variability in the estimates included in the meta-analysis, the authors suggest that the range to use for sensitivity analysis

should be based on raw study median values, which range from $3.7 million to $8.1 million. The value of $6 million for VSL equates to an average VSLY of $252 014 ($155 409 to $340 219), using a discount rate of 3 per cent over an estimated 40 years of remaining life expectancy. The meta-analysis defines cost–benefit analysis as the net present value (NPV) of dollar costs compared with the NPV of benefits, and cost–efficacy analysis as the dollar costs per outcome measure, such as dollars per life years saved.

The Council of Australian Governments (COAG) guidelines require a cost–benefit analysis as part of the impact statement for making or varying a legal instrument (e.g. a NEPM). The approach to assessing the costs associated with mortality should therefore provide an estimate of the NPV of dollar costs compared with the NPV of benefits. This is consistent with the approach outlined in this document for estimating the number of deaths avoided associated with a range of potential air quality standards. The costs of avoided deaths should be based on the VSL identified in the meta-analysis, using the average value of $6 million. Sensitivity analysis should be conducted using the recommended range of values (i.e. $3.7 million to $8.1 million).

The VSLY should be used in cost-efficacy analysis and with the recommended values: that is, an average VSLY of $252 014 ($155 409 to $340 219), using a discount rate of 3 per cent over an estimated 40 years of remaining life expectancy.

The Access Economics report (2008) also discusses, along with other approaches, the potential application of DALYs and QALYs in cost–benefit analysis and cost–utility analysis, and in assessing the benefits of an intervention.

Significant work has been undertaken in Australia in recent years to develop a burden-of-disease approach using DALYs as a measure of risk. The concept of DALYs has been introduced into the *Australian Guidelines for Water Recycling: Managing Health and Environmental Risks – Phase 1* and is being used in the second phase of these guidelines. The revision of the Australian drinking water guidelines will align with this newer approach to risk. This is an option that requires further work in relation to air pollution, and may become a valuable tool in the future to assist in analysis of costs and benefits.

For analyses of other health outcomes, such as hospital admissions and emergency department attendances, data on the costs of these outcomes are available from jurisdictional health departments or AIHW.

For outcomes such as work-loss days, data on the costs associated with loss of productivity should be calculated from the average daily wage plus on-costs as a proxy for a work-loss day. These data are available from the ABS and are updated regularly.

Exposure to air pollution is associated with increased use of medication (e.g. asthma inhalers). However, costing the use of such medication is difficult because the actual usage data are difficult to obtain; for example, because asthma relievers, such as ventolin, are available without prescription. If the costs of medication usage attributable to air pollution need to be calculated, it will be necessary to use prescription data available through the Pharmaceutical Benefits Scheme (PBS) or AIHW, but this will be a significant underestimate of the total usage. The costs of this medication are also available through the PBS or AIHW.

The costs associated with GP visits for exacerbation of existing diseases such as asthma, COPD or ischemic heart disease can be obtained from the jurisdictional health agencies or AIHW. Data from AIHW is usually a national estimate. Costs of medical treatment —through either hospitals or GPs — may vary between jurisdictions. If a national estimate of cases of a health outcome attributable to air pollution is made, then it is appropriate to use the national costing from AIHW. If avoided

health outcomes are analysed on a jurisdictional level, then data specific for the jurisdiction may be required and should be obtained from the jurisdictional health agency.

Where health outcomes are not be able to be costed, an estimate of the number of people affected should be included. Where Australian data on health costs are not available, international costings can be used, but this should be clearly identified in the analysis.

A recent study[6] commissioned by the Australian Government as a source of information on cost-benefit analysis may be useful. The study reviews European, United States and Australian Cost-Benefit Analysis (CBA) literature related to ambient air pollution, followed by a critique of the CBA methodology. It focuses on the methodology to assess the economic cost of the human health effects of ambient air pollution and identifies and describes in detail the uncertainties, contentious issues and knowledge gaps associated with the CBA methodology. The report also makes recommendations based on how to address the uncertainties, contentious issues and knowledge gaps.

[6] B Jalaludin, G Salkeld, G Morgan, T Beer, YB Nisar (2009). *A Methodology for Cost-Benefit Analysis of Ambient Air Pollution Health Impacts* (at http://www.environment.gov.au/atmosphere/airquality/publications/pubs/cost-benefit-analysis.pdf)
Commonwealth Dept of Health and Ageing and enHealth Council, 2003, Guidelines *for Economic Evaluation of environmental Health and Assessment.* http://enHealth.nphp.gov.au/council/pubs/pdf/eee guides1.pdf

8 WHERE TO FROM HERE

This paper discusses how health hazards and exposure to air pollutants are assessed, and how risk is characterised. Its purpose is to encourage discussion on the issues, processes and recommendations put forward. The aim is to establish an agreed approach that the NEPC can use to set air quality standards. The feedback provided will help to ensure that the process for setting standards and its outcomes are transparent.

8.1 THE NEXT STEPS

This paper is available on the EPHC website[7] for comment until 1 April 2010. All submissions will be regarded as public documents unless clearly marked 'confidential' and they may be made available to other interested parties, subject to provisions of the *Freedom of Information Act 1982*. No formal response will be provided on submissions. In the light of the comments received, a method will be developed and presented to EPHC for consideration.

8.2 FORM OF SUBMISSION

An electronic form for lodging comments is available. To receive the form, contact the NEPC Service Corporation or download it from the EPHC website. The form can be filled out and submitted electronically.

Hardcopy submissions should be unbound, so that they can easily be photocopied. Electronic submissions should preferably be provided as a Microsoft Word if possible. Submissions can also be made by:

- email to: kscott@ephc.gov.au)
- CD Rom
- mail to: Ms Kerry Scott
  Project Manager
  NEPC Service Corporation
  Level 5/81 Flinders Street
  ADELAIDE SA 5000

  Fax (08) 8224 0912

Submissions should be received by the NEPC Service Corporation by close of business 1April 2010**.**

[7] http://www.ephc.gov.au

APPENDIX 1: SUGGESTED TERMS OF REFERENCE FOR AN EXPERT PANEL

An expert group can be constituted in the form used by the US EPA or the WHO and United Kingdom. The actual model adopted for NEPM work will be decided by the NEPC, based on the scope of the required work. For example, a new NEPM or a full review of NEPM standards may require the WHO and United Kingdom approach, whereas the approach adopted by the US EPA may be sufficient for a less complex review.

The group should include experts from relevant scientific disciplines, preferably with expertise in air quality hazard assessment. The group will provide expertise for the extensive tasks ahead and a forum for the discussions necessary for high-quality scientific judgement on air quality standard matters relevant for Australia.

The suggested terms of reference for the expert group are:

- To conduct and/or review the hazard assessment as described in this document, including exposure response relationship, using the best available scientific evidence on the air pollutants in question.
- To comment on the adequacy of the hazard assessment underlying current standards and therefore the need for revision, or the need for development of a new standard.
- To provide its advice to the review team in a format that allows the NEPC to effectively build hazard assessment into the process for NEPM development and consideration.

In meeting the terms of reference, the expert group would be required to undertake the following tasks, in accordance with the brief provided by AAQ NEPM project team:

- Appraise existing hazard assessments for currency and adequacy, including relevant international reviews.
- Develop search criteria and strategies for selection of relevant studies published since available reviews were conducted.
- Identify and select relevant studies.
- Undertake a systematic review or report using weight of evidence methodology (as set out in the following sections).
- Provide a transparently written conclusion on the best available evidence, and recommendations that indicate whether standards should be revised.
- Recommend suitable exposure–response relationships for quantification of the health effects.
- Recommend uncertainty factors.

Adherence to the procedures described above will avoid issues concerning lack of transparency attributed to the work of an expert group.

APPENDIX 2: RISK CHARACTERISATION IN AUSTRALIA: A CASE STUDY $PM_{2.5}$

The approach taken by the US EPA was used in developing the advisory reporting standards for $PM_{2.5}$ in the AAQ NEPM. The approach is consistent with the steps outlined in Section 5.1. Air quality data from Brisbane, Melbourne, Perth and Sydney used as the basis of the exposure assessment. A set of potential standards, daily and annual, was identified, and the distribution of air pollution in each city was adjusted to simulate the distribution if pollution levels did not exceed these standards. The risk was characterised for each city individually, and the results were also combined to give an overall risk estimate for the Australian population represented by these cities.

The health outcomes considered as the basis of the risk calculations were daily mortality (all cause, respiratory and cardiovascular), hospital admission (COPD, asthma and cardiovascular disease) and annual mortality (all cause, lung cancer and cardiopulmonary disease). The risk estimates were presented as attributable cases for each of the potential standards considered. The susceptible groups considered were the entire population of these cities, children and older adults. Exposure response functions were adopted from studies conducted in the United States, because the Australian data available at the time were insufficient.

To determine the risk associated with levels of $PM_{2.5}$, data on the population were combined with the dose–response data, baseline health-incidence data and exposure data. This step was taken for all the scenarios outlined below, and for all health endpoints identified as appropriate for the basis of the standards. Baseline mortality data were obtained from the ABS, and hospital admissions data from the respective state health departments.

To assess potential costs and benefits associated with 24-hour standards for $PM_{2.5}$, a range of scenarios were proposed:

- current air quality
- reductions in $PM_{2.5}$ levels such that peak levels did not exceed 35 $\mu g/m^3$
- reductions in $PM_{2.5}$ levels such that peak levels did not exceed 30 $\mu g/m^3$
- reductions in $PM_{2.5}$ levels such that peak levels did not exceed 25 $\mu g/m^3$
- reductions in $PM_{2.5}$ levels such that peak levels did not exceed 20 $\mu g/m^3$.

To assess potential annual standards for $PM_{2.5}$, a range of annual average scenarios were assessed:

- current air quality
- reductions in $PM_{2.5}$ levels such that annual averages did not exceed 10 $\mu g/m^3$
- reductions in $PM_{2.5}$ levels such that annual averages did not exceed 8 $\mu g/m^3$
- reductions in $PM_{2.5}$ levels such that annual averages did not exceed 5 $\mu g/m^3$.

Table 9.2 summarises the outcomes of the risk characterisation process for the base case; that is, the number of health outcomes attributable to current levels of $PM_{2.5}$ in each city. The data presented are the combined data for all cities, compiled from the results of the analysis for individual cities. The risk estimates quoted relate only to anthropogenic sources of $PM_{2.5}$. Background levels have been estimated as the lowest fifth percentile of the monitored levels, consistent with the approach taken by the US EPA.

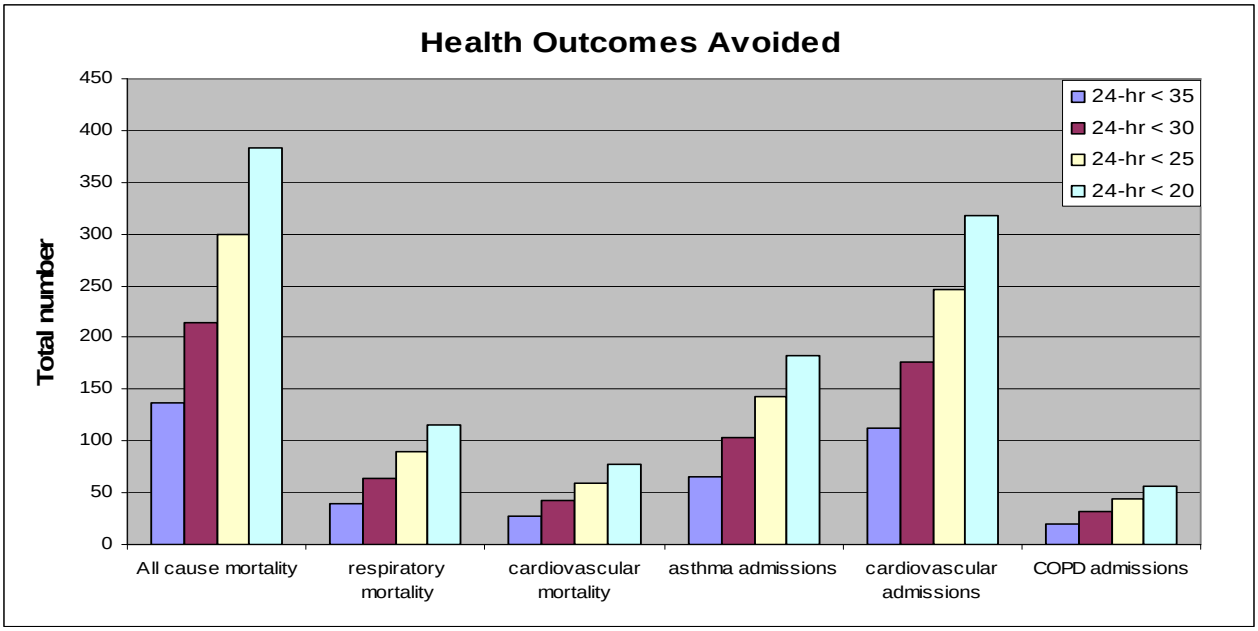
Table A 2.1 Health effects attributable to current levels of $PM_{2.5}$
(in Brisbane, Melbourne, Perth and Sydney)

	Short-term health endpoint						Long-term health endpoint		
	Mortality			Hospital admissions			Mortality		
	All cause	Respiratory	Cardio-vascular	Asthma	Cardio-vascular disease	COPD	All cause	Lung cancer	Cardio-pulmonary disease
Sydney	274	81	55	157	246	58	699	88	527
Melbourne	207	60	41	78	157	15	524	58	316
Brisbane	97	32	20	37	63	10	226	26	143
Perth	52	19	10	27	50	10	142	20	97
TOTAL	**632**	**193**	**127**	**302**	**523**	**94**	**1611**	**195**	**1096**
Including 2001 major bushfires									
Sydney	290	85	58	167	262	61	743	93	560
Brisbane	99	33	21	41	71	11	252	29	160

The health outcomes avoided (combined analysis) by reductions in $PM_{2.5}$ levels to meet each of the scenarios are shown in Figures 9.1 and 9.2. The results presented show the number of health outcomes avoided relative to the base case (i.e. current ambient levels of $PM_{2.5}$ experienced in the four cities).

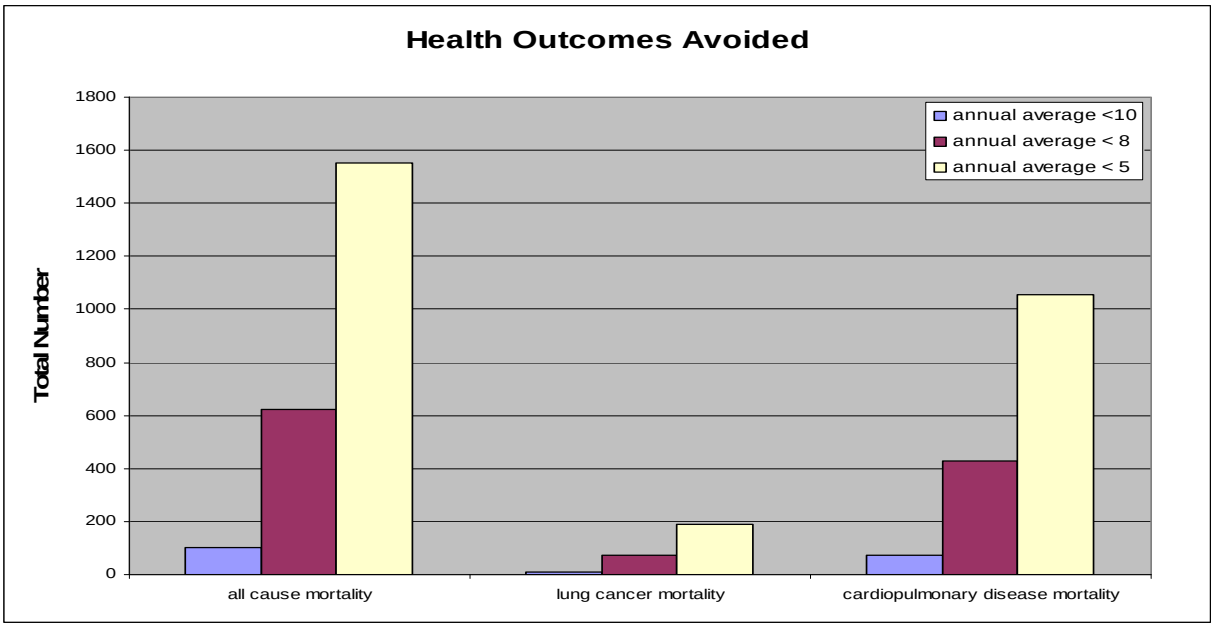
The data in Figure 9.1 show that, for each adverse health outcome, reducing $PM_{2.5}$ levels leads to significant savings in terms of adverse health effects avoided. The number of the health effects avoided depends on the extent of the reduction achieved; for example, 120 premature deaths are avoided from all cause mortality if $PM_{2.5}$ levels are reduced such that peak levels do not exceed 35 $\mu g/m^3$. This increases to about 350 premature deaths avoided if peak levels are reduced to less than 20 $\mu g/m^3$.

Figure A 2.1 Short-term health outcomes avoided



The data presented in Figure 9.2 show that, for each long-term adverse health effect, the impact of reducing $PM_{2.5}$ levels is greater than observed for short-term outcomes.

Figure A 2.2 Long-term health outcomes avoided



These data show that reductions in annual average levels of $PM_{2.5}$ are related to greater savings in adverse health outcomes than are associated with short-term exposures in both absolute and relative terms. Improving air quality to achieve an annual average of 10 μg/m³ does not provide significant health savings because most cities already meet this level. However, if levels are reduced to meet an annual average of 8 μg/m³, then for all cause mortality, the prediction is that 582 premature deaths are avoided. Most cities are close to or already meet this level; therefore, this is a realistic target and would provide significant health protection if attained.

An annual average of 5 μg/m³ is close to the estimated natural background levels of $PM_{2.5}$ and is an unrealistic target, unlikely to be achievable in any jurisdiction.

Risk estimates are associated with a significant amount of uncertainty, which arises from various sources. Some of the key uncertainties in the risk analysis for $PM_{2.5}$ included:

- the function used to model the dose–response function, which may not be the best model of the true dose–response function
- transferability of dose–response functions from overseas to the Australian context
- extrapolation of the dose–response relationships beyond the concentrations used in the epidemiological study from which they were derived
- adequacy of the air monitoring data in estimating population exposure to $PM_{2,5}$
- consideration of particle mass but not of particle composition
- error involved in the monitoring methods used for measuring $PM_{2.5}$
- assumptions used in the roll-back procedure, which may not reflect the actual distribution if $PM_{2.5}$ levels were reduced
- the estimate of background concentrations, which may not be accurate.

The overall effect of these uncertainties in the estimate of risk is unknown — it can lead to either overestimation or underestimation of the predicted effects. However, because only a subset of health effects that have been associated with exposure to $PM_{2.5}$ in overseas studies have been assessed for the purposes of this variation, the health effects presented in this report must be considered as an underestimate of the true health effects associated with $PM_{2.5}$.

Potential health costs avoided

The conclusion drawn for the scenarios described above related to the populations in Brisbane, Melbourne, Perth and Sydney statistical divisions. If the scenarios were projected for the total Australian population (2001 census 19.4 million), the health incidence savings could be approximately doubled, because the population in the four cities above account for 52 per cent of the total population and 50 per cent of the population aged 65 and over. This is considered to be a reasonable approximation, given the expected wide range in air quality between regional areas — that is, some regional areas would have significantly better air quality than urban areas, whereas others would be worse (e.g. those affected by wood smoke).

Source documents for disease category costings and the subsequent assessment of savings in terms of health costs associated with hospitalisations are from the AIHW (Mathers et al 1986, 1999). Health-related costs per case arising from hospitalisation were assessed for asthma (ninth revision of the International Classification of Disease [ICD 9] code 493), cardiovascular disease (ICD9 codes 390-398, 401-405, 410-417, 420-429), and COPD (ICD9 codes 490-492, 494, 496).Costs apportioned to hospital admissions for each disease classification include average hospital admission, medical, pharmaceutical and allied health services costs, based on the method of Mathers et al (1998a).

To project average costings (accounting for inflation) for health services from the data published by AIHW in 1993–94, the 'total health price index' was used (AIHW 2001). The averaged costs for hospital admissions were derived for asthma ($8875), cardiovascular disease ($11,709) and COPD ($9610).

If the costs of these disease states are multiplied by the projected annual health incidence savings in each scenario and then summed, the total annual monetary value for avoided health costs can be calculated. These estimates are shown in Table 9.2

Table A 2.2 Estimated short-term health costs avoided (morbidity)

24-hour scenario $PM_{2.5}$ (μg/m³)	Avoided costs of hospital admissions ($) per year			Total savings ($)
	Asthma	Cardiovascular disease	Chronic obstructive pulmonary disease	
< 35	585,750	1,323,117	192,200	2,101,067
< 30	914,125	2,072,493	307,520	3,294,138
< 25	1,269,125	2,892,123	422,840	4,584,088
< 20	1,624,125	3,711,753	547,770	5,883,648

To account for total annual health savings on top of the monetary savings for hospitalisations for the 24-hour scenarios, preventable deaths due to $PM_{2.5}$ exposure need to be included. Mortality due to respiratory and cardiovascular disease, and for all cause mortality were considered to be the appropriate indicators to be modelled. These estimates are shown in Table 9.3.

Table A 2.3 Estimated short-term health effects avoided (mortality)

24-hour scenario $PM_{2.5}$ (μg/m³)	Short-term health endpoint mortality causes		
	Respiratory	Cardiovascular	All cause
< 35	40	28	137
< 30	64	43	214
< 25	89	60	299
< 20	115	77	383

For the long-term standard, health endpoints of mortality due to lung cancer, cardio-pulmonary disease and for all causes were considered as the appropriate indicators. These were modelled for three scenarios of annual $PM_{2.5}$ levels, shown in Table 9.4.

Table A 2.4 Estimated long term health effects avoided (mortality)

Annual scenario $PM_{2.5}$ (µg/m³)	Long term health endpoint mortality causes		
	Lung cancer	Cardiopulmonary disease	All cause
< 10	12	75	100
< 8	74	428	624
< 5	188	1056	1552

Estimates of the lives that could, potentially, be saved were not given monetary values due to difficulties in estimation and in coming to a consensus on the methods for such estimations (e.g. number of years of life saved, the potential earning capacity of the individual during the years saved, and willingness and capacity to pay to save a life). However, reductions in short-term (24 hour) and annual concentrations of $PM_{2.5}$ can save lives and lead to significant monetary savings through avoiding hospitalisations. Costs due to restricted activity days and productivity losses were not estimated.

APPENDIX 3: APPLICATION OF A PRECAUTIONARY APPROACH INTERNATIONALLY

In developing air quality standards, there is always uncertainty in the use of data derived from:

- toxicological studies and controlled exposure studies, where animals or people (usually healthy or with mild symptoms of disease) are exposed to a single pollutant
- epidemiological studies, where whole populations are exposed to a pollutant mix and health effects are attributed to specific pollutants within that mix.

Added to this complexity is the transferability of the results of epidemiological studies to other populations or countries; for example, are data derived from epidemiological studies conducted in the northern hemisphere applicable in Australia?

In the United States, the US EPA and Cal EPA base their air quality standards on scientific evidence of health effects attributable to air pollution. These organisations do not explicitly consider the precautionary principle when determining the final numerical value of the standard.

The United Kingdom develops air quality standards and objectives by considering the weight of evidence from epidemiological and toxicological studies about the health effects of air pollution, and does apply the precautionary principle when determining the final numerical value of the standard. For example, controlled exposure and toxicological studies indicate that there might be a threshold for the effects of nitrogen dioxide. In contrast, the epidemiological studies show that health effects occur below this level and that there is no identifiable threshold for effect. Taking a precautionary approach, the United Kingdom applies additional safety factors in developing the standards, to account for the results of epidemiological studies showing no threshold for effects.

The WHO has published several documents about the application of the precautionary principle in developing environmental policy (WHO 2004ab). At the Fourth Ministerial Conference on Environment and Health, held in Budapest in 2004, the WHO proposed an approach for applying the precautionary principle to decisions aimed at protecting the health of children from environmental hazards (WHO 2004a). This approach is targeted at not only dealing with immediate risks to children's health but also to risks that may occur in the future. The approach is based on applying simple steps, scientific research and policy actions, as outlined below:

1. Determining whether an uncertain risk or problem merits a more thorough review — whether there is sufficient evidence to indicate a potential problem, or whether the cost of review is disproportionate to the cost of considered actions, including inaction. Sometimes a screening process may be useful.
2. Broadly defining problems to capture root sources of risks, where appropriate.
3. Considering and examining all available relevant evidence on exposure, hazard and risk in an interdisciplinary manner, and taking account of variability as well as relevant direct, indirect, cumulative and interactive effects. This can include conducting routine health and environmental monitoring to provide a baseline understanding of health and ecological impacts, as well as health trends.
4. Considering the application of simplifying rules of thumb, safety factors, default values or proxy indicators of exposure and effects when information is lacking.
5. Comprehensively examining uncertainty and gaps in information, performing sensitivity analyses, and identifying research and other ways to reduce uncertainties and gaps in knowledge, where appropriate.
6. Examining a wide range of options to reduce risks, as well as considering their trade-offs, advantages and disadvantages.
7. Determining an appropriate course of action based on scientific evidence, the examination of alternatives and public input. This involves considering a wide variety of policy tools to

implement preventive or protective actions, along with their economic, technical and political feasibility.

8. Instituting post-implementation measures, to ensure continuous risk reduction, and to understand the positive and negative impacts of interventions, and possible unintended consequences. This involves evaluating measures taken and not taken, to minimise unexpected adverse impacts and to maximise learning.

Many of these issues have been identified and included in the hazard and exposure assessment sections of this document for use in setting air quality standards in Australia.

The WHO further note that, under the approach outlined above for application of the precautionary principle, that there is no single recipe for taking precaution. What is considered an 'acceptable risk' or sufficient evidence to act is a function not only of the level of risk and the strength of evidence and uncertainty, but also the magnitude, reversibility and distribution of the risk; the availability of opportunities to prevent risk; the public's risk aversion; and society's culture and values. Decisions made using this approach should be based on the best available evidence, in addition to informed judgement and common sense. Rigorous, high-quality science, which is explicit about its limitations and gaps, is critical in the application of precaution to the protection of health.

APPENDIX 4: INTERNATIONAL APPROACHES TO ENVIRONMENTAL AND SOCIAL EQUITY

United States

In the United States, environmental and social equity issues (also referred to as social and environmental *justice* issues) are considered when making policy decisions and managing air quality. The US EPA notes that environmental justice means not only protecting human health and the environment for everyone, but also ensuring that all people are treated fairly and are given the opportunity to participate meaningfully in the development, implementation and enforcement of environmental laws, regulations and policies. In keeping with that definition, the National Environmental Justice Advisory Committee [8] (NEJAC) distinguishes public participation from meaningful involvement. The committee defines 'meaningful involvement' as requiring fully informed participants from all perspectives, all armed with the capacity to participate in policy development as full partners (NEJAC 2006). The point of meaningful involvement is not only to formulate policy advice; it must also lead the way to concrete results in achieving fair treatment. NEJAC's definition of meaningful involvement is consistent with the recommendations made by the RATF for community involvement in the development of air quality standards (NEPC 2001).

Extensive consultation with vulnerable populations — including low socioeconomic and ethnic groups within the population — informs and empowers communities to make decisions about actions that affect their health. Consultative forums are held in areas where environmental justice issues exist. With respect to setting air quality standards, where information on exposure–response functions is available for these groups, that information is considered in the risk assessment process. Where information is missing, research has been initiated, including community-based health research.

The US EPA's eight national environmental justice priorities are:

- reduced asthma attacks
- reduced exposure to air toxics
- compliance ensured
- reduced incidence of elevated blood lead levels
- fish and shellfish safe to eat
- water safe to drink
- revitalisation of brown-fields and contaminated sites
- collaborative problem solving.

California

The Cal EPA environmental justice programs focus on developing and conducting public health and environmental protection programs, policies and activities in a manner that promotes equity and affords fair treatment, accessibility and protection for all Californians, regardless of race, age, culture, income or geographic location.

Cal EPA is required to implement an environmental justice action plan to:

- develop guidance on precautionary approaches
- develop guidance on cumulative impacts analysis
- improve tools for public participation and community capacity building
- ensure environmental justice considerations within the Governor of California's environmental action plan.

[8] NEJAC is a Federal Advisory Committee to the US EPA on Environmental Justice Issues. It is a public advisory committee that provides independent advice and recommendations on the issue of environmental justice to the Administrator and other officials of the US EPA.

The environmental justice action plan must be implemented with meaningful stakeholder input. The plan is intended to assess different environmental scenarios, identify challenges and opportunities, explore the practical application of strategies, and develop recommendations to address environmental justice issues.

Two discussion papers have been released for public consultation dealing with precautionary approaches and cumulative impacts.[9]

The Cal EPA is also developing public participation guidelines, and protocols to resolve environmental justice complaints. The Cal EPA environmental justice web page is regularly updated to provide information and tools to facilitate public participation and community capacity building.

The environmental action plan outlines initiatives to protect and restore California's air, water and landscapes. Several of these initiatives have significant environmental justice implications in relation to air quality, including a focus on protecting children's health and reducing air pollution by 50 per cent. The action to take gross-polluting vehicles off the road has significant environmental justice implications, because many people in low socioeconomic groups live near major roadways.

European Union and United Kingdom

The European Union has adopted the Aarhus Convention, which was established by the United Nations in 2001. The convention builds on the Stockholm Convention and the Rio Declaration. The United Kingdom ratified the Aarhus Convention in February 2005, and it came into force in that country in May 2005.

The convention itself has three strands, designed to:

- enable people to obtain all the information they need about environmental matters, quickly and easily
- involve the public in shaping decisions about the environment, creating a partnership between citizens and public authorities
- offer people access to justice if they need to challenge environmental decisions.

In implementing the Aarhus Convention, the European Union has developed or is in the process of developing directives addressing each of these three strands.[10]

The Aarhus Convention sets out the fundamental right 'to freedom, equality and adequate conditions of life, in an environment of a quality that permits a life of dignity and well-being', by referring to principle 1 of the Stockholm Declaration. It also recalls principle 10 of the Rio Declaration, which brings in the aspect of public participation in environmental issues.

The Aarhus Convention adopts a rights-based approach, and requires parties to guarantee rights of access to information, public participation in decision making and access to justice in environmental matters. It also refers to the goal of protecting the right of every person of present and future generations to live in an environment adequate to health and wellbeing. It further

[9] These papers are available at http://www.oehha.ca.gov

[10] Directive 2003/35/EC providing for public participation in respect of the drawing up of certain plans and programmes relating to the environment was adopted on 25 June 2003. The Directive is predominantly a technical measure. It amends public participation rights in the Environmental Impact Assessment (EIA) Directive (85/337/EEC) and the Integrated Pollution Prevention and Control (IPPC) Directive (96/61/EC). It also lays down rules for public participation in plans and programmes drawn up within other existing Directives including the 1996 ambient air quality Directive (96/62/EC). The requirements of the Aarhus Convention have been incorporated into the draft EU Directive for Air Quality released in 2005.

develops the concepts of fundamental rights in the field of the environment, and places these in the context of human health and sustainable development. This structure recognises that public participation is a critical tool in guaranteeing the right to a healthy environment, environmental rights and the role of public participation in the context of sustainable development. The Aarhus Convention establishes minimum standards to be achieved, but does not stop any party adopting measures that go further in providing access to information, public participation or justice. It prohibits discrimination on the basis of citizenship, nationality or domicile against persons, seeking to exercise their rights under the convention.

ACRONYMS AND ABBREVIATIONS

AAQ	Ambient Air Quality
ABS	Australian Bureau of Statistics
AIHW	Australian Institute of Health and Welfare
APEX model	Air Pollution Exposure model
ATS	American Thoracic Society
BMC	bench mark concentration
BMCL	bench mark concentration limit
Cal EPA	California Environmental Protection Agency
COPD	chronic obstructive pulmonary disease
CSAF	chemical specific adjustment factor
DALY	disability-adjusted life year
DEFRA	Department for Environment, Food and Rural Affairs (United Kingdom)
enHealth	Environmental Health Committee
EPHC	Environment Protection and Heritage Council
GP	general practitioner
HAPEM	Hazardous Air Pollution Exposure Model
ICD	International Classification of Diseases
IGAE	Intergovernmental Agreement on the Environment
LOAEL	lowest-observed-adverse-effect level
NAAQS	National Ambient Air Quality Standards
NEPC	National Environment Protection Council
NEPC Act	*National Environment Protection Council Act 1994*
NEPM	National Environment Protection Measure
NHMRC	National Health and Medical Research Council
NOAEL	no-observed-adverse-effect level
NGO	nongovernment organisation
OEHHA	Office for Health Hazard Assessment, Californian Environmental Protection Agency
PM	particulate matter
PM_{10}	particles of 10 micrometres or less
PM_{25}	particles of 25 micrometres or less
QALY	quality-adjusted life year
RATF	Risk Assessment Task Force
RAWG	Risk Assessment Working Group
RCTs	randomised control trials
RfC	reference concentration
SSWG	Standard Setting Working Group
TEOM	Tapered Element Oscillating Microbalance
US EPA	United States Environmental Protection Agency
WHO	World Health Organization

GLOSSARY

These terms are taken from the International Programme on Chemical Safety Harmonisation Project.

Term	Description
Acceptable daily intake	Estimated maximum amount of an agent, expressed on a body mass basis, to which individuals in a (sub) population may be exposed daily over their lifetime without appreciable health risk. Related terms: *Reference dose, Tolerable daily intake*
Acceptable risk	A risk management term. The acceptability of the risk depends on scientific data; social, economic and political factors; and the perceived benefits arising from exposure to an agent.
Adverse effect	Change in the morphology, physiology, growth, development, reproduction, or life span of an organism, system, 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.
Analysis	Detailed examination of anything complex, made in order to understand its nature or to determine its essential features.
Assessment	Evaluation or appraisal of an analysis of facts and the inference of possible consequences concerning a particular object or process.
Assessment end-point	Quantitative or qualitative expression of a specific factor with which a risk may be associated, as determined through an appropriate risk assessment.
Assessment factor	Numerical adjustment used to extrapolate from experimentally determined (dose–response) relationships to estimate the agent exposure below which an adverse effect is unlikely to occur. Related terms: *Safety factor, Uncertainty factor*
Concentration	Amount of a material or agent dissolved or contained in unit quantity in a given medium or system.
Concentration–effect relationship	Relationship between the exposure, expressed in concentration, of a given organism, system or (sub)population to an agent in a specific pattern during a given time and the magnitude of a continuously graded effect to that organism, system or (sub)population. Related terms: *Effect assessment, Dose–response relationship*
Dose 1	Total amount of an agent administered to, taken up by, or absorbed by an organism, system or (sub) population.
Dose–effect relationship	Relationship between the total amount of an agent administered to, taken up by, or absorbed by an organism, system or (sub)population and the magnitude of a continuously graded effect to that organism, system or (sub)population. Related terms: *Dose–response relationship, Effect assessment, Concentration–effect relationship*
Dose-related effect	Any effect to an organism, system or (sub)population as a result of the quantity of an agent administered to, taken up by, or absorbed by that organism, system or (sub)population.

Term	Description
Dose–response	Relationship between the amount of an agent 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. Synonymous with: *Dose–response relationshi*p Related terms: *Dose–effect relationship, Effect assessment, Concentration–effect relationship*
Dose–response assessment	Analysis of the relationship between the total amount of an agent administered to, taken up by, or absorbed by an organism, system or (sub)population and the changes developed in that organism, system or (sub)population in reaction to that agent, and inferences derived from such an analysis with respect to the entire population. Dose–response assessment is the second of four steps in risk assessment. Related terms: *Hazard characterisation, Dose–effect relationship, Effect assessment, Dose–response relationship, Concentration–effect relationship*
Dose–response curve	Graphical presentation of a dose–response relationship.
Dose–response relationship	Relationship between the amount of an agent 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. Related terms: *Dose–effect relationship, Effect assessment, Concentration–effect relationship*
Effect	Change in the state or dynamics of an organism, system or (sub)population caused by the exposure to an agent.
Effect assessment	Combination of analysis and inference of possible consequences of the exposure to a particular agent based on knowledge of the dose–effect relationship associated with that agent in a specific target organism, system or (sub)population.
Expert judgment	Opinion of an authoritative person on a particular subject.
Exposure 1	Concentration or amount of a particular agent that reaches a target organism, system or (sub)population in a specific frequency for a defined duration.
Exposure assessment 1	Evaluation of the exposure of an organism, system or (sub)population to an agent (and its derivatives). Exposure assessment is the third step in the process of risk assessment.
Exposure scenario 1	A set of conditions or assumptions about sources, exposure pathways, amounts or concentrations of agent(s) involved, and exposed organism, system or (sub)population (e.g. numbers, characteristics and habits) used to aid in the evaluation and quantification of exposure(s) in a given situation.
Fate	Pattern of distribution of an agent, its derivatives, or metabolites in an organism, system, compartment or (sub)population of concern as a result of transport, partitioning, transformation or degradation.
Guidance value	Value, such as concentration in air or water, that is derived after allocation of the reference dose among the different possible media (routes) of exposure. The aim of the guidance value is to provide quantitative information from risk assessment to the risk managers to enable them to make decisions. Related term: *Reference dos*e

Term	Description
Hazard	Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub)population is exposed to that agent.
Hazard assessment	A process designed to determine the possible adverse effects of an agent or situation to which an organism, system or (sub)population could be exposed. The process includes hazard identification and hazard characterisation. It focuses on the hazard, in contrast to risk assessment, where exposure assessment is a distinct additional step.
Hazard characterisation	The qualitative and, where possible, quantitative description of the inherent property of an agent or situation having the potential to cause adverse effects. Where possible, it should include a dose–response assessment and its attendant uncertainties. Hazard characterisation is the second stage in the process of hazard assessment and the second of the four steps in risk assessment. Related terms: *Dose–effect relationship, Effect assessment, Dose–response relationship, Concentration–effect relationship*
Hazard identification	The identification of the type and nature of adverse effects that an agent has an inherent capacity to cause in an organism, system or (sub)population. Hazard identification is the first stage in hazard assessment and the first of the four steps in risk assessment.
Margin of exposure	Ratio of the no-observed-adverse-effect level (NOAEL) for the critical effect to the theoretical, predicted, or estimated exposure dose or concentration. Related term: *Margin of safety*
Margin of safety	For some experts, margin of safety has the same meaning as margin of exposure, while for others margin of safety means the margin between the reference dose and the actual exposure. Related term: *Margin of exposure*
Measurement endpoint	Measurable (ecological) characteristic that is related to the valued characteristic chosen as an assessment point.
Pathophysiology	Deranged function in an individual or an organ that is due to a disease. A pathophysiologic alteration is a change in function as distinguished from a structural defect.
Reference dose	An estimate of the daily exposure dose that is likely to be without deleterious effect even if continued exposure occurs over a lifetime. Related term: *Acceptable daily intake*
Response	Change in the state or dynamics of an organism, system or (sub)population in reaction to exposure to an agent.
Risk	The probability of an adverse effect in an organism, system or (sub)population caused under specified circumstances by exposure to an agent.
Risk analysis	A process for controlling situations where an organism, system or (sub)population could be exposed to a hazard. The risk analysis process consists of three components — risk assessment, risk management and risk communication.

Term	Description
Risk assessment	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 agent, taking into account the inherent characteristics of the agent of concern as well as the characteristics of the specific target system. The risk assessment is the first component in a risk analysis process and includes four steps — hazard identification, hazard characterisation (related term: *Dose–response assessmen*t), exposure assessment and risk characterisation.
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 characterisation is the fourth step in the risk assessment process.
Risk communication	Interactive exchange of information about (health or environmental) risks among risk assessors, managers, news media, interested groups and the general public.
Risk estimation	Quantification of the probability, including attendant uncertainties, that specific adverse effects will occur in an organism, system or (sub)population due to actual or predicted exposure.
Risk evaluation	Establishment of a qualitative or quantitative relationship between risks and benefits of exposure to an agent, involving the complex process of determining the significance of the identified hazards and estimated risks to the system concerned or affected by the exposure, as well as the significance of the benefits brought about by the agent. Risk evaluation is an element of risk management and is synonymous with risk–benefit evaluation.
Risk management	Decision-making process involving considerations of political, social, economic and technical factors with relevant risk assessment information relating to a hazard, used to develop, analyse and compare regulatory and non-regulatory options, and to select and implement appropriate regulatory response to that hazard. Risk management comprises three elements — risk evaluation, emission and exposure control, and risk monitoring.
Risk monitoring	Process of following up the decisions and actions within risk management in order to ascertain that risk containment or reduction with respect to a particular hazard is assured. Risk monitoring is an element of risk management.
Safety	Practical certainty that adverse effects will not result from exposure to an agent under defined circumstances. It is the reciprocal of risk.
Safety 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. Related terms: *Assessment factor, Uncertainty factor*
Threshold	Dose or exposure concentration of an agent below which a stated effect is not observed or expected to occur.
Tolerable daily intake	Analogous to *Acceptable daily intak*e. The term 'tolerable' is used for agents that are not deliberately added, such as contaminants in food.

Term	Description
Tolerable intake	Estimated maximum amount of an agent, expressed on a body mass basis, to which each individual in a (sub)population may be exposed over a specified period without appreciable risk.
Toxicity	Inherent property of an agent to cause an adverse biological effect.
Uncertainty	Imperfect knowledge about the present or future state of an organism, system or (sub)population under consideration.
Uncertainty factor	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. Related terms: *Assessment factor, Safety factor*
Validation	Process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose. 'Reliability' is defined as the reproducibility of outcome of the approach, method, process or assessment over time. 'Relevance' is defined as the meaningfulness and usefulness of the approach, method, process or assessment for the defined purpose.

[1] This term is also contained in the list of IPCS key exposure assessment terminology — both definitions are consistent and interchangeable, depending on user preference.

REFERENCES AND BIBLIOGRAPHY

Abt Associates (2005). *Particulate Matter Health Risk Assessment for Selected Urban Areas*. Prepared for the Office of Air Quality Planning and Standards, US Environmental Protection Agency, Research Triangle Park, NC. http://www.epa.gov/ttn/naaqs/standards/pm/data/pm_risk_tsd_finalreport_2005_mainbody.pdf

Access Economics (2008). The Health of Nations: The Value of a Statistical Life, Report prepared for Australian Safety and Compensation Council

AIHW (Australian Institute of Health and Welfare) (2001). *Australia's Health Services Expenditure to 1999–00*, Health Expenditure Bulletin No. 17, Health and Welfare Expenditure Series No. 12, AIHW, Canberra.

AIHW (Australian Institute of Health and Welfare) (2006). The burden of disease and injury in Australia 2003, available http://www.aihw.gov.au/publications/hwe/bodaiia03/bodaiia03.pdf

AIHW (Australian Institute of Health and Welfare) (2007). *The Burden of Disease and Injury in Australia 2003*, PHE 82, AIHW, Canberra. http://www.aihw.gov.au/publications/index.cfm/title/10317

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 Science USA* 70(8):2281–2285.

ATS (American Thoracic Society) (2000). What constitutes an adverse health effect of air pollution? Official statement of the American Thoracic Society. *American Journal of Respiratory and Critical Care Medicine* 161(2:Pt 1):t-73.

Barnett AG, Williams GM, Schwartz J, Best TL, Neller AH, Petroeschevsky AL and Simpson RW (2006). The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australian and New Zealand cities. *Environmental Health Perspectives* 114(7):1018–1023.

Barnett AG, Williams GM, Schwartz J, Neller AH, Best TL, Petroeschevsky AL and Simpson RW (2005). Air pollution and child respiratory health: a case-crossover study in Australia and New Zealand. *American Journal of Respiratory and Critical Care Medicine* 171(11):1272–1278.

Bell ML, Dominici F and Samet JM (2005). A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. [see comment]. *Epidemiology* 16(4):436–445.

Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T and Friedenreich C (1999). Traditional reviews, meta-analyses and pooled analyses in epidemiology. *International Journal of Epidemiology* 28:1–9.

Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D and Farland W (2006). *Critical Reviews in Toxicology* 36(10):781–792.

Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J, and Vickers C (2008). *Critical Reviews in Toxicology* 38(2):87–96.

Brauer M, Blair J and Vedal S (1996). Effect of ambient ozone exposure on lung function in farm workers. *American Journal of Respiratory and Critical Care Medicine* 154(4:Pt 1):t-7.

BTRS (Bureau of Transport and Regional Services) (2005). *Health Impacts of Transport Emissions in Australia: Economic Costs*, BTRS, Canberra.

CDC (Center for Disease Control and Prevention) (2005). *Third National Report on Human Exposure to Environmental Chemicals,* National Center for Environmental Health, Atlanta, GA. http://www.cdc.gov/ExposureReport/pdf/thirdreport.pdf

Clancy L, Goodman P, Sinclair H and Dockery DW (2002). Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. *Lancet,* 360(9341):1210–1214.

Cohen AJ, Anderson HR, Ostro B, Pandey KD, Krzyzanowski M, Künzli N, Gutschmidt K, Pope CA, Romieu I, Samet JM and Smith KR (2004). Urban air pollution. In: *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*, vol 2, Ezzati M, Lopez AD, Rodgers A and Murray CJL (eds) World Health Organization, Geneva, 1353–1433.

Commonwealth Dept of Health and Ageing and enHealth Council, 2003, Guidelines *for Economic Evaluation of environmental Health and Assessment.* http://enHealth.nphp.gov.au/council/pubs/pdf/eee guides1.pdf

COT (Committee on toxicity of chemicals in food, consumer products and the environment) (2007). Statement on the COT workshop on evolving approaches to chemical risk assessment. COT Statement March 2007.

Covello VT and Merkhofer MW (1993). *Risk Assessment Methods – Approaches for Assessing Health and Environmental Risks*, Plenum Press, New York

Crump KS (1984). A new method for determining allowable daily intakes. *Fundamental and Applied Toxicology* 4(5):854–871.

de Hollander AEM, Melse JM, Lebret E and Kramers PGN (1999). An aggregate public health indicator to represent the impact of multiple environmental exposures. *Epidemiology* 10:606–617.

Desqueyroux H, Pujet JC, Prosper M, Le Moullec Y and Momas I (2002b). Effects of air pollution on adults with chronic obstructive pulmonary disease. *Archives of Environmental Health* 57(6):554–560.

Desqueyroux H, Pujet JC, Prosper M, Squinazi F and Momas I (2002a). Short-term effects of low-level air pollution on respiratory health of adults suffering from moderate to severe asthma. *Environmental Research* 89(1):29–37.

Diaz E, Smith-Sivertsen T, Pope D, Lie RT, Diaz A, McCracken J, Arana B, Smith KR and Bruce N (2007). Eye discomfort, headache and back pain among Mayan Guatemalan women taking part in a randomised stove intervention trial. *Journal of Epidemiology and Community Health* 61(1):74–79.

Dockery DW, Luttmann-Gibson H, Rich DQ, Link MS, Mittleman MA, Gold DR, Koutrakis P, Schwartz JD and Verrier RL (2005). Association of air pollution with increased incidence of ventricular tachyarrhythmias recorded by implanted cardioverter defibrillators. *Environmental Health Perspectives* 113(6):670–674.

Doll R and Peto R (1981). The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *Journal of the National Cancer Institute* 66:1193–1308.

Donaldson K, Brown D, Clouter A, Duffin R, MacNee W, Renwick L, Tran L and Stone V (2002). The pulmonary toxicology of ultrafine particles. *Journal of Aerosol Medicine* 15(2):213–220.

Donaldson K, Gilmour MI and MacNee W (2000). Asthma and PM10. *Respiratory Research* 1(1):12–15.

Dybing E and Soderlund EJ (1999). Situations with enhanced chemical risks due to toxicokinetic and toxicodynamic factors. *Regulatory Toxicology and Pharmacology* 30(2):S27–30.

Dybing E and Totlandsdal AI (2004). *Air Pollution and the Risks to Human Health – A Toxicological Perspective*. http://airnet.iras.uu.nl (accessed 12 March 2009)

enHealth (2004). *Environmental Health Risk Assessment: Guidelines for Assessing Human Health Risks from Environmental Hazards*, Australian Government Department of Health and Ageing, Canberra.
http://www.health.gov.au/pubhlth/strateg/envhlth/risk (accessed 14 June 2006)

EPHC (Environment Protection and Heritage Council) (2004). *Time Activity Study, Summary of Findings*. Research report to the Environment Protection and Heritage Council by the Centre for Population Studies in Epidemiology, South Australian Department of Human Services. http://www.ephc.gov.au/taxonomy/term/21

EPHC (Environment Protection and Heritage Council) (2005). *National Environment Protection (Air Toxics) Measure. Air Toxics Tier 2 Prioritisation Methodology*, EPHC.

Ezzati M, Hoorn SV, Rodgers A, Lopez AD, Mathers CD and Murray CJL (2004). Potential health gains from reducing multiple risk factors. In: *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*, vol 2, Ezzati M, Lopez AD, Rodgers A, and Murray CJL (eds), World Health Organization, Geneva, 2167–2190

Ezzati M, Lopez AD, Rodgers A and Murray CJL (eds) (2004). *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*, World Health Organization, Geneva, 1229–1230.[

Filipsson AF, Sand S, Nilsson J and Victorin K (2003). The benchmark dose method—review of available models, and recommendations for application in health risk assessment. *Critical Reviews in Toxicology* 33(5):505–542.

Filliger, P, Puybonnieux-Texier V and Schneider J (1999). PM_{10} population exposure. Health costs due to road traffic-related air pollution. An impact assessment project of Austria, France and Switzerland. Prepared for the Third WHO Ministerial Conference on Environment and Health, London, 16–18 June 1999.

Gielen MH, van der Zee SC, van Wijnen JH, van Steen CJ and Brunekreef B (1997). Acute effects of summer air pollution on respiratory health of asthmatic children. *American Journal of Respiratory and Critical Care Medicine* 155(6):2105–2108.

Gilboa SM, Mendola P, Olshan AF, Langlois PH, Savitz DA, Loomis D, Herring AH and Fixler DE (2005). Relation between ambient air quality and selected birth defects, seven county study, Texas 1997–2000. *American Journal of Epidemiology* 162(3):238–252.

Gold DR, Damokosh AI, Pope CA, Dockery DW, McDonnell WF, Serrano P, Retama A and Castillejos M (1999). Particulate and ozone pollutant effects on the respiratory function of children in southwest Mexico City. *Epidemiology* 10(1):8–16. (erratum in 10:470)

Greenland S and Robins JM (1998). Conceptual problems in the definition and interpretation of attributable fractions. *American Journal of Epidemiology* 128(6):1185–1197.

Guyatt GH, Berman LB, Townsend M, Pugsley SO and Chambers LW (1987). A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 42(10):773–778.

Hedley AJ, Wong C, Thuan Qouc Thach, Stefan MA, Lam T and Anderson HR (2002). Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: an intervention study. *Lancet* 360(9346):1646–1652.

Hertel RF (1996). Outline on risk assessment of existing substances in the European Union. *Environmental Toxicology and Pharmacology* 2:93–96.

Hill AB (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine* 58:295–300.

Hiltermann TJ, Stolk J, van der Zee SC, Brunekreef B, de Bruijne CR, Fischer PH, Hiemstra PS and van Bree L (1998). Asthma severity and susceptibility to air pollution. *European Respiratory Journal* 11(3):686–693.

Holguín F, Téllez-Rojo MM, Hernández M, Cortez M, Chow JC, Watson JG, Mannino D and Romieu I (2003). Air pollution and heart rate variability among the elderly in Mexico City. *Epidemiology* 14(5):521–527.

IGAE (Intergovernmental Agreement on the Environment) (1992). *National Environment Protection Council Act 1994* No. 126, 1994 — Schedule 1.

IPCS (International Programme on Chemical Safety) (1994). *Environmental Health Criteria 170: Assessing Human Health Risks of Chemicals: Derivation of Guidance Values for Health-Based Exposure Limits*, World Health Organization, Geneva. http://www.inchem.org/documents/ehc/ehc/ehc170.htm

IPCS (International Programme on Chemical Safety) (2004). *Principles for Modelling Dose-Response for the Risk Assessment of Chemicals,* draft, World Health Organization, Geneva. http://www.who.int/ipcs/methods/harmonization/draft_document_for_comment.pdf (accessed June 2006)

IPCS (International Programme on Chemical Safety) (2005). *Chemical-Specific Adjustment Factors for Interspecies Differences and Human Variability: Guidance Document for use of Data in Dose/Concentration-Response Assessment*. Harmonization Project Document No. 2. International Programme on Chemical Safety, World Health Organization, Geneva.

Jalaludin B, G Salkeld, G Morgan, T Beer, YB Nisar (2009). *A Methodology for Cost-Benefit Analysis of Ambient Air Pollution Health Impacts* (at http://www.environment.gov.au/atmosphere/airquality/publications/pubs/cost-benefit-analysis.pdf)

Janes HL, Sheppard L and Lumley T (2005). Case-crossover analyses of air pollution exposure data: referent selection strategies and their implications for bias. *Epidemiology* 16:717–726.

Johnston FH, Bailie RS, Pilotto LS and Hanigan IC (2007). Ambient biomass smoke and cardio-respiratory hospital admissions in Darwin, Australia. *BMC Public Health* 7:240 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2220001

Jones PW (1993). *Measurement of Health-related Quality of Life in Asthma and Chronic Airways Disease*, Kluwer Academic Publishers, Dordrecht, Boston, 301–320.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A and Anderson HR (1997). Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. Air pollution and health: a European approach. *BMJ* 314(7095):1658–1663.

Kinney PL, Thurston GD and Raizenne M (1996). The effects of ambient ozone on lung function in children: a reanalysis of six summer camp studies. *Environmental Health Perspectives* 104(2):170–174.

Korrick SA, Neas LM, Dockery DW, Gold DR, Allen GA, Hill LB, Kimball KD, Rosner BA and Speizer FE (1998). Effects of ozone and other pollutants on the pulmonary function of adult hikers. *Environmental Health Perspectives* 106(2):93–99.

Künzli N, Kaiser R, Medina S, Studnicka M, Chanel O and Filliger P (2000). Public health impact of outdoor and traffic-related air pollution. *Lancet* 356:795–801.

Künzli N, Kaiser R, Medina S, Studnicka M, Oberfeld G and Horak F Jr (1999). Air pollution attributable cases. Technical report on epidemiology. Health costs due to road related air pollution. An impact assessment project of Austria, France and Switzerland. Prepared for the Third Ministerial Conference on Environment and Health, London, 1999. Berne, Switzerland, Federal Department for Environment, Transport, Energy and Communication. Bureau for Transport Studies, Switzerland.

Laden F, Schwartz J, Speizer FE and Dockery D (2006). Reduction in fine particulate air pollution. *American Journal of Respiratory and Critical Care Medicine* 173(6):667–672

Lambert WE, Samet JM, Stidley CA, and Spengler JD (1992). Classification of residential exposure to nitrogen dioxide. *Atmospheric Environment* 26A:2185–2192

Lippmann M and Schlesinger RB (2000). Toxicological bases for the setting of health-related air pollution standards. *Annual Review of Public Health* 21:309–333.

Lucas RM and McMichael AJ (2005). Association or causation: evaluating links between 'environment and disease'. *Bulletin of the World Health Organization* 83(10):792–795.

Marks GB, Dunn SM and Woolcock AJ (1992). A scale for the measurement of quality of life in adults with asthma. *Journal of Clinical Epidemiology* 45(5):461–472.

Mathers C and Penm R, 1999, *Health System costs of cardiovascular diseases and diabetes in Australia 1993 1999.*

Mathers C et al, 1998a, *Disease costing methodology used in the disease costs and impact study 1993-1994*

Mathers C et al, 1998b, *Health System costs of Diseases and Injury in Australia 1993-1994*

MRC (Medical Research Council, Great Britain) (1999). *Risk Assessment Approaches used by the United Kingdom Government for Evaluating Human Health Effects of Chemicals*, Institute for Environment and Health, Leicester, UK.

Muller P (2002). *Potential for Occupational and Environmental Exposure to Ten Carcinogens in Toronto*. Prepared for Toronto Public Health, ToxProbe Inc. http://www.toronto.ca/health/pdf/cr_technicalreport.pdf

Murray CJL and Lopez AD (eds) (1996). *The Global Burden of Disease; a Comprehensive Assessment of Mortality and Disability from Disease, Injury, and Risk Factors in 1990 and Projected to 2020*, Harvard University Press, Cambridge, MA. (Global burden of disease and injury series, Vol. 1)

Navidi W (1998). Bidirectional case-crossover designs for exposures with time trends. *Biometrics* 54(2):596–605.

NEJAC 2006, Future Mechanisms To Enhance Stakeholder Involvement And Engagement To Address Environmental Justice , Report to USEPA, available at:
http://www.epa.gov/compliance/resources/publications/ej/nejac/stakeholder-involv-9-27-06.pdf

NEPC (National Environment Protection Council) (2007). Ambient air quality, National Environment Protection Measure. Review discussion paper. NEPC, Adelaide.

NHMRC (National Health and Medical Research Council) (2006). *Ambient Air Quality Standards Setting: An Approach to Health-Based Hazard Assessment*, National Health and Medical Research Council and Environmental Health Committee.

OEHHA (Office of Environmental Health Hazard Assessment) (1999). *Air Toxics Hot Spots Program Risk Assessment Guidelines. Part 1: The Determination of Acute Reference Exposure Levels for Airborne Toxicants,* OEHHA, California Environmental Protection Agency, California, 33.

OEHHA (Office of Environmental Health Hazard Assessment) (2000). *Air Toxics Hot Spots Program Risk Assessment Guidelines. Part 3: The Determination of Chronic Reference Exposure Levels for Airborne Toxicants,* OEHHA, California Environmental Protection Agency, California.

Pilotto LS, Nitschke M, Smith BJ, Pisaniello DL, Ruffin RE, McElroy HJ, Martin J and Hiller JE (2004). Randomized controlled trial of unflued gas heater replacement on respiratory health of asthmatic schoolchildren. *International Journal of Epidemiology* 33(1):208–211

Pope CA, Burnett RT, Thun MJ, Calle EE, Kreski D, Ito K and Thurston GD (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association* 287:1132–1141.

RATF (Risk Assessment Taskforce) (2000). *Report of the Risk Assessment Taskforce.* Prepared for the National Environment Protection Council, National Environment Protection Council, Adelaide. http://www.nepc.gov.au/sites/default/files/AAQ_RATF_Rpt__RATF_Report_200010.pdf

RAWG (Risk Assessment Working Group) (2002). *Report of the Risk Assessment Working Group.* Prepared for the National Environment Protection Council, National Environment Protection Council, Adelaide.

Rich DQ, Mittleman MA, Link MS, Schwartz J, Luttmann-Gibson H, Catalano PJ, Speizer FE, Gold DR and Dockery DW (2006). Increased risk of paroxysmal atrial fibrillation episodes associated with acute increases in ambient air pollution. *Environmental Health Perspectives* 114(1):120–123.

Ritz B, Wilhelm M, Hoggatt K and Ghosh G (2007). Ambient air pollution and preterm birth in the environment and pregnancy outcomes study at the University of California, Los Angeles. *American Journal of Epidemiology* 166(9):1045–1052.

Rockhill B, Newman B and Weinberg C (1998). Use and misuse of population attributable fractions *American Journal of Public Health* 88(1):15–19, (erratum in 98(12):2119

Roemer W, Clench-Aas J, Englert N, Hoek G, Katsouyanni K, Pekkanen J and Brunekreef B (1999). Inhomogeneity in response to air pollution in European children (PEACE project). *Occupational and Environmental Medicine* 56(2):86–92.

Romieu I, Meneses F, Ruiz S, Sienra JJ, Huerta J, White MC and Etzel RA (1996). Effects of air pollution on the respiratory health of asthmatic children living in Mexico City. *American Journal of Respiratory and Critical Care Medicine* 154(2:Pt 1):t-7.

Ross MA, Persky VW, Scheff PA, Chung J, Curtis L, Ramakrishnan V, Wadden RA and Hryhorczuk DO (2002). Effect of ozone and aeroallergens on the respiratory health of asthmatics. *Archives of Environmental Health* 57(6):568–578.

Rothman KJ and Greenland S (1998). *Modern Epidemiology,* 2nd edition, Lippincott-Raven, Philadelphia.

Rothman KJ and Greenland S (2005). Causation and causal inference in epidemiology. *American Journal of Public Health* 95(S1):S144–S150.

Salem H and Katz SA. (2006) Inhalation toxicology (illustrated). Edition: 2, CRC Press, 2006

Samet J and Jaakkola JJ (1999). The epidemiological approach to investigating outdoor air pollution. In: *Air Pollution and Health,* Holgate ST, Samet J, Koren HS and Maynard RL (eds), Academic Press, London, 431–460.

Schwartz J (2000). Harvesting and long term exposure effects in the relation between air pollution and mortality. *American Journal of Epidemiology* 151(5):440–448.

Seethaler R (1999). Health costs due to road traffic-related air pollution. An impact assessment project of Austria, France and Switzerland. Prepared for the WHO Ministerial Conference on Environment and Health, London, June 1999. Synthesis report prepared on behalf of the tri-lateral research team of Austria, France and Switzerland.

Simpson R, Williams G, Petroeschevsky A, Best T, Morgan G, Denison L, Hinwood A, Neville G and Neller A (2005). The short-term effects of air pollution on daily mortality in four Australian cities. *Australian and New Zealand Journal of Public Health* 29(3):205–212.

Simpson R, Williams G, Petroeschevsky A, Best T, Morgan G, Denison L, Hinwood A and Neville G (2005). The short-term effects of air pollution on hospital admissions in four Australian cities. *Australian and New Zealand Journal of Public Health* 29(3):213–221.

Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, Grant D, Hartley M, Knaap A, Kroese D, Mangelsdorf I, Meek E, Rice JM and Younes M (2001). IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regulatory Toxicology and Pharmacology* 34(2):146–152

Travis KZ, Pate I and Welsh ZK (2005). The role of the benchmark dose in a regulatory context. *Regulatory Toxicology and Pharmacology* 43:280–291.

Tunnicliffe WS, Burge PS and Ayres JG (1994). Effect of domestic concentrations of nitrogen dioxide on airway responses to inhaled allergen in asthmatic patients. *Lancet* 344(8939–8940):1733–1736.

US EPA (United States Environmental Protection Agency) (1986). *Guidelines for Carcinogen Risk Assessment*, EPA/630/P-03/001B, US EPA, Washington, DC.

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

US EPA (United States Environmental Protection Agency) (1996a). *Air Quality Criteria for Ozone and Related Photochemical Oxidants*, EPA/600/p93/004a-cF, US EPA, Washington, DC.

US EPA (United States Environmental Protection Agency) (1996b). *Air Quality Criteria for Particulate Matter,* NC EPA 600/P-95/001a-cF, US EPA, Research Triangle Park, NC.

US EPA (United States Environmental Protection Agency) (2003). *Benchmark Dose Software (BMDS) Tutorial.* http://www.epa.gov/ncea/bmds_training

US EPA (United States Environmental Protection Agency) (2004). *Air Quality Criteria for Particulate Matter*, final report, EPA 600/P-99/002aF-bF, US EPA, Washington, DC.

US EPA (United States Environmental Protection Agency) (2005a). *Guidelines for Carcinogen Risk Assessment*, EPA/630/P-03/001B, US EPA, Washington, DC.

US EPA (United States Environmental Protection Agency) (2005b). *Air Quality Criteria for Ozone and Related Photochemical Oxidants,* first external review draft, EPA/600/R-05/004aA-cA, US EPA, Washington, DC,

US EPA (United States Environmental Protection Agency) (2006). *Air Quality Criteria for Ozone and Related Photochemical Oxidants*, final, EPA/600/R-05/004aF-cF, US EPA, Washington, DC

US EPA (United States Environmental Protection Agency) (2008). *Integrated Science Assessment for Oxides of Nitrogen - Health Criteria,* EPA/600/R-08/071, US EPA, Research Triangle Park, NC. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=194645 (accessed September 2008)

USEPA 2008, Integrated Scientific Assessment for Sulfur Oxides – Health Criteria, September 2008, EPA/600/R-08/047F

Vedal S (1997). Health effects of ambient particles: lines that divide. *Journal of the Air and Waste Management Association* 47:551–581.

Vedal S, Petkau J, White R and Blair J (1998). Acute effects of ambient inhalable particles in asthmatic and nonasthmatic children. *American Journal of Respiratory and Critical Care Medicine* 157(4:Pt 1):t-43.

Ware JE Jr and Sherbourne CD (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 30(6):473–483

WHO (World Health Organization) (2000). *Air Quality Guidelines for Europe*, 2nd edition, WHO regional publications, European series, No. 91, WHO Regional Office for Europe, Copenhagen.

WHO (World Health Organization) (2003). *Health Aspects of Air Pollution: Particulate Matter, Ozone and Nitrogen Dioxide*, WHO Regional Office for Europe, Copenhagen

WHO (World Health Organization) (2004). *Health Aspects of Air Pollution: Results from the WHO Project 'Systematic Review of Health Aspects of Air Pollution in Europe'*, WHO Regional Office for Europe, Copenhagen.

WHO (World Health Organization) (2004). *Principles for Modelling Dose-Response for the Risk Assessment of Chemicals*, WHO, Geneva.

WHO (World Health Organization) (2005a). *Air Quality Guidelines for Europe,* 2nd edition, WHO European Series No. 91, WHO Regional Office for Europe, Copenhagen.

WHO (World Health Organization) (2005b). *Chemical-Specific Adjustment Factors for Interspecies Differences and Human Variability: Guidance Document for Use of Data in Dose/Concentration-Response Assessment*, Harmonization Project Document No. 2, WHO, Geneva.

WHO (World Health Organization) (2006). *Air Quality Guidelines. Global update 2005. Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide,* WHO, Geneva.

WHO (World Health Organization) (2006). *Principles for Evaluating Health Risks in Children Associated with Exposure to Chemicals. Environmental Health Criteria 237,* WHO, Geneva.

WHO (World Health Organization) Working Group (2000). Evaluation and use of epidemiological evidence for environmental health risk assessment: WHO guideline document. *Environmental Health Perspectives* 108(10):997–1002.