PROTOCOL
FOR THE DEVELOPMENT OF NATIONAL AMBIENT AIR QUALITY OBJECTIVES

Part 1

SCIENCE ASSESSMENT DOCUMENT AND DERIVATION OF THE REFERENCE LEVEL(S)

November 1996

Prepared by the

Federal–Provincial Working Group on

Air Quality Objectives and Guidelines

This report is published under section 8 of the Canadian Environmental Protection Act (CEPA).

Any inquiries concerning this publication or requests for copies should be directed to either of the following officials:

Director
Science Assessment and
Policy Integration Division
Atmospheric Environment Service
Environment Canada
4905 Dufferin Street
Toronto, Ontario M3H 5T4

Director
Bureau of Chemical Hazards
Environmental Health Centre
Health Canada
Tunney's Pasture
Ottawa, Ontario K1A 0L2

Final technical editing and desktop publishing by Astroff Corkum Ross Associates Inc.

November 1996
ISBN 0-662-25637-9
Catalogue En42-17/5-1-1997E

Over 50% recycled paper
including 10% post-
consumer fibre.

M—An official mark of Environment Canada

TABLE OF CONTENTS

LIST OF ACRONYMS

CEPA	Canadian Environmental Protection Act
FPAC	Federal–Provincial Advisory Committee
FEV1	Functional Exhale Volume
GLP	Good Laboratory Practices
IUPAC	International Union for Pure and Applied Chemistry
LED	Lower confidence limit on an Effective Dose
LOAEL	Lowest Observed Adverse Effect Level
LOEL	Lowest Observed Effect Level
MAC	Management Advisory Committee
NAAQO	National Ambient Air Quality Objective
NAICC	National Air Issues Coordinating Committee
NOAEL	No Observed Adverse Effect Level
NOEL	No Observed Effect Level
PAH	Poly Aromatic Hydrocarbons
PAN	Peroxyacetyl Nitrate
PBPK	Physiological Based Pharmacokinetic
PEM	Personal Exposure Modelling
pNEM	Probabilistic National Air Quality Standards Exposure Model
RfD	Reference Dose
SAR	Structure-Activity Relationships
SI	International System of Units
U.S. EPA	United States Environmental Protection Agency
WGAQOG	Working Group on Air Quality Objectives and Guidelines

1 INTRODUCTION

National Ambient Air Quality Objectives (NAAQO) are benchmark levels of protection for people and the environment in Canada. The criteria for their development and use must be based on scientifically defensible evidence while incorporating a margin of protection that reflects three factors: variability in the levels of exposure and their associated effects; uncertainty in the analysis, and the values of Canadian society.

This document explains how scientific information is obtained, evaluated, and published in support of NAAQOs. The rest of this section outlines the legal basis and development process. Users of this document include evaluators, contractors or others who perform work for the Working Group on Air Quality Objectives and Guidelines (WGAQOG), as well as risk managers, members of industry and the general public. Publication of these guidelines makes the information on the principles, concepts and methods used available to other agencies, provinces, industry, academia, and interested members of the public.

In the objectives development process, there may be intrinsic tensions among desirable goals that are difficult to accomplish together. One such challenge is the balance between the need for a firm scientific foundation for setting ambient air quality objectives and the need for protection in the face of uncertainty and incomplete information. Another is the need for effective protection of human and environmental receptors and the difficulty in achieving technical and social change in a short time period. A challenge also exists in setting objectives that apply to all parts of the country in an even and fair manner while recognizing the diversity that exists and the need to adapt recommendations to local needs. The process described in this protocol cannot resolve the tensions associated with these challenges; it can, however, make them explicit and facilitate their resolution through the risk management process.

1.1 LEGAL BASIS

The Canadian Environmental Protection Act (CEPA), passed into law in 1988, establishes a comprehensive legal framework for the management of toxic substances in Canada. It addresses pollution problems on land, in water, and through all layers of the atmosphere.

NAAQOs are developed by the WGAQOG, which reports to the Federal–Provincial Advisory Committee (FPAC) under CEPA. Representatives of federal, provincial and territorial departments of environment and health form this Working Group.

Provincial governments have the primary responsibility in many areas of air pollution control, with federal actions integrated with those of the provinces. The objectives are published under CEPA, and the provinces may adopt them through processes of their choice.

1.2 GENERAL CONCEPTS

The five guiding principles used by the WGAQOG in developing NAAQOs, as

approved by CEPA/FPAC in November 1994 are:

I. be consistent with the philosophy of CEPA;

II. recognize the variable sensitivities of subgroups of the Canadian population and of particular ecosystems and organisms in the environment. Given the large range of these sensitivities, it may not be possible to protect every sensitive individual and ecosystem from all effects;

III. provide a range of levels reflecting the range of biological responses and sensitivities, allowing for various regulatory options to accommodate regional priorities, while endeavouring to maintain consistent national levels of environmental quality;

IV. be reasonable, workable and useable, reflecting a consultative process that includes government, industry, public advocacy groups and the Canadian public and recognizes the importance of scientific, social and economic considerations; and

V. be based on recognized scientific principles and include risk assessment and risk management. The scientific basis for objectives should be presented in a manner which is readily accessible to, and which can be understood by, the Canadian public.

Further, the WGAQOG also applies the following concepts, such that NAAQOs will:

I. contribute to sustainable development through pollution prevention, the ecosystem approach, maintenance of biodiversity, and the precautionary principle;

II. consider other sources of exposure on a chemical-specific basis to account for the quality and quantity of total exposure; and

III. follow a development and consultation process that is fully transparent.

1.3 NAAQO DEVELOPMENT PROCESS

The goal of this protocol is to promote the consistency and scientific quality of NAAQOs. The process is sequential. A description of the major stages in the process follows, together with identifying also the section in this protocol which addresses each stage.

1. Identify a Need (section 2)

Substances may be proposed for the development of air quality objectives by the WGAQOG membership, CEPA/FPAC or the National Air Issues Coordinating Committee (NAICC). The WGAQOG will determine if it is appropriate to proceed with the development of an air quality objective, based upon guidelines laid out in the air quality objective derivation protocol.

2. Scientific Assessment and Evaluation (sections 3 through 13)

Scientific review, assessment and evaluation will be undertaken by external or in-house federal or provincial resources. Recently prepared federal or provincial documents will be utilized to the fullest extent possible, as will documents from other agencies and countries. The WGAQOG will review the currently available scientific information on human health, vegetation, animal, material and aesthetic impacts to identify dose-response relationships for a variety of

receptor endpoints (e.g., crop yield, respiratory irritation, mortality).

3. Derivation of the Reference Level (section 14)

This section includes a summary of all of the scientific data compiled above, then, through appropriate statistical techniques or other analyses, the Reference Level is determined. This is a level above which there are demonstrated effects on human health and/or the environment. It provides a scientific basis for establishing goals for air quality management. The Reference Level may be proposed for one or more time periods (e.g., 24 hours, 30 days), and one or more receptors. This number is based on an interpretation of the best available scientific evidence.

4. Reviews (section 15)

Federal and provincial members of WGAQOG review the document during development. External scientific peer reviews are commissioned as appropriate. There may be limited provincial consultations, primarily with government stakeholders (e.g., program/policy managers) at this time.

5. Tabling of the Science Assessment Document with CEPA/FPAC

The Scientific Review Document (sections A through N) will be tabled with CEPA/FPAC. It is anticipated that some general discussion will occur at this stage to develop options for the recommended form and level of the air quality objective(s).

Development of National Ambient Air Quality Objective Options (Protocol, Part II)

The WGAQOG will determine the likelihood of impact and develop options for air quality objectives. This discussion will be elaborated on in Part II. Recommendations for the NAAQO will be based upon a risk-assessment approach, and will be designed to provide national guidance for protection of the general public and the environment. Each province has the option to use the Scientific Review Document (plus any other

products of this process) and initiate public consultations leading to province-specific air quality objectives.

6. Air Quality Objectives Options Development

A Rationale Document summarizes the relevant scientific information, risk assessment, and current and anticipated exposures. It presents the rationale for selecting the form and level of the NAAQO and identifies the form(s) and level(s) of the recommended NAAQO. The WGAQOG presents the recommended objectives to CEPA/FPAC and the NAICC.

7. WGAQOG Revisits NAAQO Recommendation

Following the decision(s) and guidance of CEPA/FPAC or the NAICC, it may be necessary for the WGAQOG to review and revise the NAAQO recommendation. Subsequent to this review, the revised recommended NAAQO would be submitted again.

8. NAAQO Proposal Package

This package contains the recommended objective(s) and includes an economic assessment (if required). The responsibility for developing this package rests with CEPA/FPAC or their designate.

9. NAAQO Approval

The proposed NAAQO is reviewed and approved by Environment Canada, Health Canada, CEPA/FPAC and CEPA/MAC (Management Advisory Committee) prior to publication in the *Canada Gazette*. The responsibility for soliciting the necessary approvals lies with CEPA/FPAC or their designate.

- Draft Order in Council;
- National Ambient Air Quality Objective(s) and the scientific Reference Level;
- Explanatory Note — reference to CEPA;
- Background Note — reference to the scientific, economic and strategic thinking which has formed the development of the Reference Level and the Air Quality Objectives.

10. NAAQO Promulgation Package and Publication in the *Canada Gazette, Part 1*

It is the responsibility of Environment Canada, Environmental Protection Service, Economic Analysis Branch, to prepare the Promulgation Package and submit it for publication in the *Canada Gazette*. The package includes:

- Recommendation to the Governor in Council, signed by the Minister of the Environment and/or the Minister of Health;

2 NOMINATION AND SELECTION OF SUBSTANCES FOR REVIEW

Nomination of substances for review are made by the public, industry, and government agencies through the following mechanisms:

- members of the Working Group on Air Quality Objectives and Guidelines;
- CEPA/Federal Provincial Advisory Committee; and
- National Air Issues Coordinating Committee.

The WGAQOG will determine the need for extensive evaluation by considering some or all of the following criteria (adapted from WHO, 1989):

- capability of the substance to cause adverse effects on human health or the environment, where irreversible effects are of special concern;
- ubiquity and abundance of the substance in the Canadian environment, particularly the atmosphere (notwithstanding its presence in the atmosphere of the built environment);
- environmental transformations to form secondary pollutants or metabolic alterations, when these alterations may lead to the production of chemicals with greater toxic potential;
- persistence in the environment, particularly if the pollutant would resist environmental degradation and accumulate in humans or food chains;
- likelihood of effects, magnitude of the population exposed, and the existence of sensitive sub-populations;
- current or potential relevance to Canada as a national concern for more than one province and/or territory, and priorities of local air quality management jurisdictions;
- appropriateness of managing the substance from a regional/airshed versus site specific approach; and
- appropriateness of managing the substance using air quality objectives, in contrast with other available or established management options (e.g., the Priority Substances List and the Strategic Options Process), recognizing that indoor air contaminants may be addressed by other mechanisms.

To support the WGAQOG deliberations, a brief review of the nominated substance may be undertaken outlining:

- key transformation reactions and major intermediate or end products;
- current and proposed emission sources in Canada (and elsewhere if the substance is subject to atmospheric long-range transport);
- ambient concentrations across Canada versus background levels (levels characteristic of locations remote from significant sources, which may reflect natural or background levels);
- temporal trends and spatial patterns;
- exposure to receptors via the atmosphere relative to other media (including indoor air quality);

- environmental persistence and potential for bioaccumulation;
- summary of effects (i.e., sensitive receptors and relevant end points);
- summary of effects of similar substances already known to cause adverse impacts;
- objectives, guidelines, standards in other jurisdictions; and
- current management strategies.

If the WGAQOG determines that an in-depth evaluation is required, then a scientific review document is prepared in accordance with this protocol. It is understood that not all data required for completion of this protocol may be available for all exposures. Thus, the protocol provides for reasonable estimates and extrapolations in a manner consistent with current practices and knowledge in risk assessment. Inevitably, however, there will be a need for scientific judgement. The purpose of this protocol is to frame the problem in such a way that uncertainty is minimized and the degree of judgement applied to the problem is both minimized and supported by the best available evidence.

3 INTERIM RECOMMENDATIONS

In the event of significant scientific uncertainty or limited information relevant to the Canadian context, the Scientific Review Document will be completed and a Reference Level identified. However, the degree of uncertainty may lead to an interim Air Quality Objective (AQO). Given the acknowledged concern over adverse effects of the air contaminant, the interim recommendation is anticipated to encourage the scientific community to address the knowledge gaps. It also signifies the need to consider a preventive, precautionary approach to the management of the air contaminant.

At the time the Reference Level is recommended, it will also be noted what additional information is required to develop a full Air Quality Objective and/or at what time interval the science behind the interim recommendation should be re-evaluated.

4 SOURCES OF INFORMATION

General guidance is provided in this section for making decisions on the acceptability of each information source. Given the diversity of sources of information, it is important to recognize that this selection process relies on expert judgement. Further, the information must be interpreted in context, with an understanding of the fundamental issues inherent in air quality studies and inhalation toxicology as they apply to the management of ambient air quality.

Only information that is publicly accessible and available for critical review can be used for the purpose of developing NAAQOs. The following sources are acceptable:

- peer reviewed scientific journal papers;
- review articles and other scientific publications (e.g., conference proceedings);
- government, industry and university reports;
- industry and trade association publications;
- reports submitted to other departments or agencies of the governments of Canada, the provinces, or the territories;
- reports or reviews prepared by other national and international agencies;
- texts and other reference books; and
- unpublished data or reports made available to the review process, provided that they can be referenced or quoted and that due weight is given to the degree of quality control exerted in data collection and the critical review they have received.

Regardless of the source, reference citations will conform to standard bibliographic practice and will allow the reader to locate and obtain a copy of the reference. Not all relevant publications are identified in electronic bibliographic databases. It will be necessary to review older journals and to contact experts in the field to identify appropriate references.

The following sources are not acceptable:

- confidential material;
- personal communications of unpublished data, opinions, etc.; and
- materials which cannot be made available to the public upon reasonable request.

5 PHYSICAL AND CHEMICAL PROPERTIES OF THE SUBSTANCE

Indicate the nature of the substance as a compound (e.g., CO), a group of compounds (e.g., dioxins), or a mixture (e.g., diesel fumes). List the IUPAC (International Union of Pure and Applied Chemistry) name, trade names and common synonyms. Describe the basic physical and chemical properties of the substance, in SI (International System of Units) units, according to the following list, as appropriate. In those cases where the pollutant is a mixture of substances, identify the major components and describe their properties as above.

Physical description (odour, colour)	Atomic number (if elemental)
Molecular formula	Atomic radius or molecular volume (if known)
Atomic weight	Molecular weight
Melting point	Molecular structure
Boiling point	Isotopes, congeners and relative abundance
Density/Relative Density	Common valences
Heat of fusion (heat of combustion)	Rate constants
Water solubility	Photolysis
Vapour pressure	Oxidation or photo-oxidation
Light absorption characteristics	Hydrolysis
CAS number	Physical state
Henry's Law Constant	Bioconcentration factor
Octanol - Water partition co-efficient	Organic – Carbon partition co-efficient

Present all physical measures in SI units; and express partition coefficients as both log and actual values, for convenience in interpretation.

6 ENVIRONMENTAL FATE AND BEHAVIOUR

Present sufficient information to adequately characterize the transformations and movement of the pollutant within the atmosphere. Where appropriate, include information to describe the cycling of the pollutant between the atmosphere, biosphere, geosphere and hydrosphere to support subsequent discussions on multi-media exposures. Present half-life information for atmospheric lifetimes (and for other relevant media) to support the discussion of transformation and long range transport.

Comparisons to natural atmospheric processes are an important part of this analysis. The relevant comparisons include both concentrations and the variability in these levels. The concentrations and variability concentration resulting from human activity are compared to levels presumed to reflect background levels in the absence of human activity. These background levels may be inferred from monitoring data in locations remote from human activity or, where known, historical levels independent of emissions from sources associated with human activity. Information on background levels may be difficult to confirm and estimation may be required.

Use sections 6.1 through 6.6 as guidance for providing the relevant information.

6.1 PARTITION PROCESSES AND TRANSFORMATION REACTIONS

Describe information on the significant partitioning processes and physical transformations and chemical (inorganic and organic) reactions in gaseous and aqueous phases and heterogeneous reactions. Take the following into consideration:

- sorption/desorption
- volatilization
- wet and dry deposition characteristics
- biotransformation
- photolysis
- oxidation/reduction
- hydrolysis

6.2 TRANSFORMATION PRODUCTS

Identify the intermediate and end products of the physical and chemical transformations. This may be done in conjunction with the review of effects to identify transformation products which may, in their own right, have adverse impacts on health or environment or exacerbate the impact of the substance under consideration.

6.3 ROLE OF METEOROLOGY

Describe the impact of meteorology (temperature, wind speed, relative humidity, atmospheric mixing, solar radiation, etc.) on the transformation reactions. This will be most relevant for secondary pollutants (e.g., ground-level ozone, formaldehyde, PAN [peroxyacetyl nitrate]).

6.4 POTENTIAL FOR ATMOSPHERIC TRANSPORT

Describe the potential for the substance to be transported through the atmosphere. This will be a function of its atmospheric persistence (as dictated by the relevant transformation reactions), meteorology and source characteristics (e.g., stack height, plume rise, turbulence at release altitude, etc.).

6.5 ATMO-GEO-BIO-HYDRO CYCLES

Where consideration of multi-media exposures is required, present information on the cycling of the substance through the ecosystem. This requires the identification of the principal media (where accumulation or exposure may be an issue) and the routes of exposure.

6.6 LOADING CAPACITIES

The loading capacity for the most sensitive relevant ecosystem will be identified and characterized to the extent known. This may involve determining the loading capacity of local soils, lakes, and airsheds, taking into account the integrity of biological communities, susceptibility of indicator species, aesthetic values, and protection of historical and cultural heritage.

7 IDENTIFICATION AND CHARACTERIZATION OF EMISSION SOURCES

Describe the anthropogenic and natural sources of the pollutant and identify their relative contributions to the atmosphere. Provide the relative proportions of total releases to the environment, to illustrate the relative loadings to different media: air, soil, and water. Describe natural sources in terms of the biotic or abiotic processes leading to their emission.

In the case of secondary pollutants, where they are not directly emitted, quantify the precursor emissions and identify reaction pathways and meteorological conditions for the generation of the secondary pollutant.

Where routine monitoring measures a surrogate for the pollutant of interest or an indicator for a class of pollutants (e.g., benzo(a)pyrene for polycyclic aromatic hydrocarbons), describe those circumstances, and quantify and describe the relationship between the measured indicator and the contaminants of toxicological interest. Describe where commonly data sets have shifted from one mode of measurement to another, or from one surrogate to another, or to a direct measurement, and quantify and describe the translation of one to the other. Where the standard technology or method of monitoring has changed, quantify and describe the translation from one to the other and specifically note the limitations of interpretation, including differences in variability using the different methods (e.g., arising from different sampling times).

Present information on anthropogenic sources to describe the direct and fugitive releases, point, area, and mobile sources. Classify the sources according to their industrial sectors and geographic distribution. For information on the temporal characteristics of the emissions, both natural and anthropogenic, provide current emission inventory data, including emission factors and historical trends. Access the most recent federal and provincial databases for this information. Where this information is insufficient, access industrial emission information through appropriate channels. Where appropriate, indicate the significance of long range transport into and within Canada.

This information is essential to the modelling of emission scenarios, the economic impact assessment activities, and stakeholder consultation activities.

8 MONITORING TECHNOLOGIES

Describe the current monitoring technologies with respect to detection limits, accuracy, precision, sampling frequency and duration, reliability, and any interferences from other substances. When more than one monitoring method is commonly used, contrast and compare the methods identifying similarities and differences in sampling characteristics. This may require the inclusion of results of laboratory instrument characterization and co-location studies. Evaluate the appropriateness of sampling frequency, duration and detection for current monitoring technologies for impact assessment. Critically evaluate direct versus indirect technologies and continuous versus grab sample or time-integrated sampling for their appropriateness in characterizing ambient concentrations and receptor impacts. Identify inadequacies in current monitoring technology and recommend future improvements according to performance specifications.

9 ENVIRONMENTAL LEVELS

The intent is to describe as completely as possible the spatial and temporal characteristics of the substance across Canada. Acquire available data from provincial, federal and municipal databases, industrial sources, and government and academic research publications, and review the process of data acquisition, transmission and quality assurance to ensure data validity and comparability.

The time period covered by the data should be sufficient to provide a representative assessment of annual trends and seasonal patterns. This time period will vary by substance as a function of the impacts of climate and economic cycles. While a five-year data record (US EPA 450/2-78-027R, 1978) would be considered sufficient to establish short- to medium-term characteristics and trends for primary air pollutants, a ten-year or longer data record is required to characterize trends in secondary pollutants such as ground-level ozone. In the trend analysis, account for the normalization for meteorological variation. In the absence of extensive ambient data sets, ambient concentrations may be estimated using atmospheric dispersion models. Use data derived from dispersion modelling only if model performance has been validated by comparison with observed concentrations under similar climate conditions and emission regimes.

Analyze and present the data to elucidate the following characteristics:

- background (i.e., natural or uncontaminated) levels;
- annual trends of parameters relevant to receptor exposure;
- seasonal patterns;
- weekly and daily patterns;
- diurnal and, where relevant (as with ozone), shorter-term patterns;
- episode characteristics;
- frequency of exceedence plots showing the existing NAAQOs or provincial standards/guidelines; and
- frequency distributions.

Present the above information for geo-political regions as defined by airshed characteristics. Identify substance concentration characteristics unique to receptor locations (e.g., rural agricultural areas, urban core, suburban residential, etc.). Present data in formats that are intuitively informative (maps, plots, tables, etc.) and that are compatible with existing reports and publications produced by federal, provincial, and territorial agencies participating in the process. These formats are necessary in order to allow for the identification of the spatial and temporal characteristics by readers familiar with such reports.

Ambient background levels are considered to have low spatial variability, and to be representative of current background levels characteristic of continental air masses that are remote from anthropogenic sources. A key characteristic of these background levels is that they show little spatial variability or long-term temporal variability compared to near-field spatial and temporal variability observed near significant emission sources.

Frequency-of-exceedence plots indicate the percentage of time a given concentration is

exceeded for a given site or region. Define these concentration increments to complement the effects information and support the exposure assessment. Where effects thresholds are not available, select uniform concentration increments to span the given concentration range. Other formats for displaying the environmental levels information will be dictated by the effects information to facilitate the exposure assessment.

10 QUALITATIVE EVALUATION OF EXPERIMENTAL STUDIES

It is essential that studies chosen for a comprehensive review and evaluation be consistently selected. Since the specific sources of information, and the amount and quality of toxicity and dose-response data, will differ both between and within the human health and the environmental effects evaluations, expert judgement will be required to determine the acceptability of a study.

No one experimental study is likely to answer all, or even more than a few, of the questions about a particular pollutant. Rather, each study can be expected to contribute a portion of the overall evaluation. On a qualitative basis, a study must meet certain basic criteria for validity, quality, and relevance in order that its conclusions can be adequately quantified and interpreted. The limitations essential to the analysis of each study must be explicitly described.

As described in the subsections below, the assessor can ask a variety of questions. However, the main issues in selecting sources are: “Is the question being posed in the study relevant?” and “Does the study design permit the investigator (either study author or assessor) to adequately quantify the answer?”

The main requirement for a qualitative assessment of the literature is expert judgement. The following elements of the scientific database (both environmental and human health effects) are not intended as an exhaustive or prescribed list of criteria. They are provided to give an assessor general guidance on the required elements leading to a qualitative evaluation of the scientific database.

10.1 CRITERIA FOR ASSESSING THE QUALITY OF THE ENVIRONMENTAL EFFECTS LITERATURE

The number of environmental receptors affected by air contaminants can be large, and include plants (indigenous and cultivated), animals (domestic and wildlife, including both vertebrates and invertebrates), microorganisms, soils, surface waters, and materials. Aesthetic considerations are also involved, and include odour, colour and visibility. Due to this diversity of receptors, a wide range of information sources must be used in evaluating the effect of an air contaminant on the environment. Expert judgement is required in all aspects of the selection process outlined below.

In addition to the general criteria given above, each source of information describing an effect or absence of effect of the air contaminant on a biological receptor should meet the following criteria:

- the test organism must be completely identified by genus and species name, and, if possible, by subspecies, cultivar, ecotype or race;
- the age or life stage of the test organism must be identified;
- genetically and phenotypically similar organisms should be used through the course of the experiments, unless comparisons of dissimilar organisms is the subject of study;
- whole organisms or excised tissues may be used (in the case of excised tissues a

complete description of the excised portion should be presented);

- if only a portion of the test organism is exposed to the pollutant, this portion should be completely identified (tissue type, age, pre-treatment conditions);
- exposure pathways or routes should be clearly identified;
- the experimental system or enclosure should be completely described (e.g., closed chamber, open-top field chamber, zonal fumigation system);
- environmental conditions (e.g., temperature, humidity, quality and quantity of light, soil characteristics, wind) during the course of each experiment should be measured and described;
- each experiment must be properly controlled (i.e., filtered and ambient air controls in experiments utilizing open-top or zonal fumigation systems may be required to account for exposure to airborne compounds not included as experimental treatments);
- each experiment must be adequately replicated;
- appropriate sampling procedures, experimental techniques and analytical technology must be used for the measurement of treatment concentrations;
- an appropriate endpoint must be measured using suitable sampling procedures, experimental techniques and analytical technology (endpoints at one or more levels of organization [biochemical, cellular, tissue, organ, whole organism] may be measured and an effect detectable at a lower level of organization should not be automatically rejected based on the absence of an observed or measurable effect at a higher level); and
- studies documenting effects of the air contaminant on biodiversity, community structure or function, or populations of species are acceptable, provided that they are completely described, and appropriate field techniques and statistical methods have been used and presented.

Each source of information describing an effect (or absence of effect) of the air contaminant on an abiotic (e.g., surface water, materials) receptor should meet the following criteria:

- the physical and chemical characteristics of the receptor must be described;
- exposure pathways or routes should be clearly identified;
- environmental conditions during the course of each experiment or survey should be measured and described;
- experiments must be properly controlled (i.e., in the case of field studies in the vicinity of point sources, an ecologically analogous site distant from the source and out of the plume may be considered an acceptable control); and
- endpoints must be completely described (e.g., erosion, chemical changes, deterioration) and measured using suitable sampling procedures, experimental techniques and analytical technology.

Aesthetic considerations are also included in the evaluation of the effect of an air contaminant on the environment. Sources of information describing the effect of an air

contaminant on odour, colour or visibility of the atmosphere should include:

- a description of the air quality in the absence of the air contaminant (specifying whether the comparison is with other near-field atmospheres that lack significant sources of the contaminant or with natural background);
- environmental conditions during the observational period (temperature, humidity, wind, light);
- a description of the sources contributing to the deterioration of the quality of the air;
- a description of the atmospheric reactions leading to the deterioration of colour or visibility, or the increase in odour of the air;
- the endpoint measured, the rationale for the choice of the endpoint, and a description of the methods of measurement;
- when the endpoint reflects aesthetic issues, there should be a full description of the aesthetic problem and the variability in perception of the problem (e.g., the range of sensitivity to low-level odours); and
- the history of the problem and trends in the parameter of interest in the area.

10.2 CRITERIA FOR ASSESSING THE QUALITY OF HUMAN HEALTH EFFECTS LITERATURE

Epidemiological and toxicological studies currently provide much of the information used in identification of human health hazards. These may require extrapolation from animals to people, an exercise somewhat arbitrary in nature. Obviously, the preferred data for human health is that derived from observations on people, provided that this meets scientific and ethical criteria. Because the human health literature is often limited in scope, subject to confounding, and lacking sensitivity and specificity, the best available alternative is the use of data from test species.

Although the use of animal data for the prediction of human health effects has limitations, the reality is that animal toxicology in the form of controlled experimental investigation has been essential for determining exposure–response relationships and for understanding the mechanisms of toxic effects. Advances in understanding the comparative pathophysiology and methods such as toxicokinetic modelling are providing a basis for translating results from animal toxicology to human risk assessment. Currently, however, application of a more or less standard set of "uncertainty factors" is used to perform these extrapolations.

10.2.1 Criteria for Selection of Surrogate Species and Systems

Given that the use of data from animal studies has its limitations, there should be an explicit rationale for selection and interpretation of animal data for human risk assessment. This rationale must address the uncertainties and limitations in the practice as applied to this evaluation, as well as the reasons for using it.

The rationale for selecting the surrogate species for *ex vivo* or *in vivo* studies, or the surrogate *in vitro* system should be clearly

stated and defended, and the component of the species or system which is serving as surrogate (e.g., whole body or respiratory tract response) should be specified.

For *ex vivo* or *in vivo* studies, the term surrogate includes laboratory rodents and non-rodents, non-human primates, (i.e. from laboratory mouse to baboon via the pig). For *in vitro* systems, the term surrogate includes cell lines and organ preparations.

10.2.2 Toxicological Studies

This is, by far, the largest source of data on surrogates for human exposure. Because of previous problems with toxicological testing, several jurisdictions have well-established protocols for test procedures, data handling and processing, (Health Canada, U.S. EPA, etc.). The following questions are relevant to the acceptance of the study under review:

- Are the characteristics well described? For example: genetic traits; sex; age and maturity; body weight; hormonal status; nutrition and diet; selection of species/strain; veterinary and immunization history; housing conditions, including lighting and caging and bedding; and employment of good laboratory practices (GLP). In the case of *in vitro* investigations, is the cell line adequately described as to source; propagation and maintenance; microbiological status; stage of cell cycle; etc.; and are the organ preparations well described as to source; maintenance, etc.?
- Are the route, duration and concentration of exposure similar to those for humans; was appropriate technology used to measure exposure concentrations for the containment and exposure of the test subjects?
- Are the anatomical, physiological, biochemical, behavioural, neurological, pathological, reproductive and other characteristics of interest relevant to those of humans?
- Are the testing methods, sources of experimental error, and changes in laboratory practices well described; and were appropriate biological and health endpoints selected?
- Is the investigational design adequate? For example, assignment to investigational groups, and difference between control groups and other groups in exposure only and not in any other aspects. What are the control groups controlling for; are confounding factors considered; are observations objective or subjective; what is the probability of replication of subjective assessments; and are there pre-determined scales for subjective assessments?
- Can the evaluator describe sources of bias in the investigation?
- Were the statistical analyses suitable for the data?
- Is this information sufficient to allow for extrapolation of *in vivo* or *in vitro* effects to humans? This may apply to the whole animal response, or to system, organ, tissue or cell responses.

10.2.3 Critical Evaluation of Epidemiological Reports

There are advantages and disadvantages to the different observational epidemiological study designs. There will be restricted

sensitivity in cases where exposure concentration is low and the results may be subject to unsuspected confounding risk factors. These should be discussed in detail and the imposed limitations accounted for. Findings from experimental epidemiological studies may be available for only a few air pollutants.

Assessment of the quality of data on health effects, exposure and interactions must consider:

- adequacy;
- reliability and robustness;
- appropriateness of findings to human exposure;
- study design strengths and weaknesses;
- validity of study design for etiologic or exposure–response investigation;
- statistical power (if negative result);
- quality of exposure assessment;
- sensitivity, specificity, and predictive value of clinical endpoints used;
- use of biomarkers appropriate to the problem under investigation;
- control of confounding factors;
- identification of bias;
- description of the population about which the question is being asked:
  - Who is included and who is excluded?
  - Are the subjects a sample of the target population?
  - How is the sample selected?
  - Is there random or systematic selection?
  - What are the possible sources of bias?
  - Is the sample large enough to answer the question(s) being addressed?

10.2.4 Predictive Modelling

The use of structure–activity relationships (SAR), physiological-based pharmacokinetic (PBPK) modelling and personal-exposure modelling (PEM) can provide useful insights and understandings. However, the link between the theory of the model and its application to the actual exposure situation must be carefully explained and justified. The following questions should be carefully evaluated:

- Can the evaluator describe the molecular basis for the proposed SAR?
- Can the evaluator describe the theoretical and mathematical bases for the PBPK model and its experimental support?
- Can the evaluator describe the theoretical and mathematical bases for the PEM model and its experimental support?

11 TOXICITY ASSESSMENT

Assessing the potential toxicity of the chemical, establishing the presence of an adverse effect, and identifying the mechanism of action, constitute the initial stages of the toxicity evaluation. Databases may vary considerably and a systematic critical qualitative and quantitative review (with appropriate level of detail for the diversity of users) and impartial presentation of all available data should be made. Recognizing that there is variation on the part of individual authors and assessors in the interpretation and extrapolation of data, the following factors may be considered when evaluating the toxicity of a pollutant to the relevant receptors. Their relevance depends upon the amount and quality of available information, and the resources available for preparing and presenting the assessment.

The qualitative and quantitative assessments form the basis of the integrated assessment. The qualitative assessment allows the assessor to gain an overall picture of the literature, and some level of confidence as to its underlying strength. For some purposes, qualitative risk assessment may be sufficient, especially if data are lacking and the evidence to support interpretation is not well structured. For most purposes, qualitative risk assessment will be a provisional step prior to quantitative risk assessment. If the level of confidence gained is sufficient, the assessor can proceed to the quantification of the adverse endpoints observed. This quantification is the last step of the toxicological assessment process.

The principal drawback of quantitative risk assessment is in the prediction of absolute risk. Quantitative risk assessment is more satisfactory for the comparison of risks than for the estimation of absolute risk. Applied to air quality objective development, this means that the estimates are generally more valid for comparisons among options for setting objectives than as a guide to the impact of any one option. The single greatest advantage of quantitative risk assessment is the rigour it imposes in guiding the review of the assumptions that underlie the analysis.

Neither qualitative nor quantitative risk assessment substitutes for judgement in the derivation of recommendations for ambient air quality objectives. However, the role of risk assessment in general is to broaden the basis of agreement, to reduce the degree of uncertainty in critical estimates, and to reduce the role of interpretation to the minimum necessary to support an informed judgement.

11.1 QUALITATIVE ASSESSMENT

The qualitative assessment establishes the adequacy of the database to derive and support the final recommendations for the specific Reference Level. In doing so, the assessor should be able to give expression to the consistency of the database, and whether or not there are irreconcilable contradictions or gaps that cannot effectively be bridged.

The data within the individual studies will often dictate the exact form of the analysis. A number of subjects must be addressed: the type of effect (threshold vs. non-threshold, acute vs. chronic, reversible vs. non-reversible); the relevance of the effect to other species or receptors; cause–effect relationships; identification of susceptible or

sensitive species, sub-groups and populations; and relevance to the Canadian context. The endpoints can be categorized in order to more exactly classify the effects observed. This list of categories, provided for guidance, is not intended to be exhaustive, or all-inclusive and they are not presented in any order of priority.

11.1.1 Human Health Effects

- increased mortality (total, respiratory, cardiovascular);
- increased incidence of cancer;
- increased hospitalization (frequency, duration and total, respiratory and cardiovascular);
- increased emergency ward, physician visits;
- increased incidence or prevalence of cough or phlegm requiring medical attention;
- increased incidence or prevalence of symptomatic asthmatic attacks;
- increased rates of disability from cardiovascular or pulmonary disease in the community;
- increasing rate of decline in pulmonary function (Functional Exhale Volume [FEV1]) relative to predicted values in adults with increasing age, or failure of children to maintain their predicted FEV1 growth curve;
- increased number of persons with FEV1 below normal limits;
- reduced ability to cope with daily activities, i.e., shortness of breath or increased anginal episodes or reduced activity days;
- increased or more intensive use of pulmonary or cardiovascular medication;
- increased incidence or prevalence of lower respiratory-tract infection;
- increased incidence or prevalence of chest tightness;
- increased incidence of upper respiratory-tract infections that interfere with normal activity;
- increased incidence or prevalence of cough or chronic cough;
- increased prevalence of wheezing in the chest apart from colds, or of wheezing most days and nights;
- headache or nausea associated with odours;
- acute upper respiratory-tract infections that do not interfere with normal activity;
- eye, nose, or throat irritation that may interfere with normal activities;
- headache or nausea, with or without other symptoms or distress, associated with odours;
- increased incidence or prevalence of chemical pneumonia and bronchitis;
- degradation of renal and neurological function;
- irritation of mucous membranes.

As indicated, this is not an exhaustive list; any endpoints which can be plausibly linked to exposure by weight of evidence or specific mechanism should be assessed.

11.1.2 Vegetation Effects

- reduced biomass yield;
- reduced fruit or seed set;

- reduced reproductive vitality;
- visual blemishes;
- increased susceptibility to pathogens;
- increased susceptibility to physical stressors (heat, moisture, etc.);
- delayed maturity;
- mortality;
- impaired photosynthetic production.

11.1.3 Animal Effects

- reduced growth and/or weight gain;
- reduced reproductive vitality;
- increased susceptibility to pathogens;
- increased susceptibility to physical stressors;
- mortality.

11.1.4 Material Effects

- increased decay rates;
- reduced structural integrity;
- discolouration/fading;
- soiling.

11.1.5 Aesthetic Effects

- sensory effects (may include odour, visibility, etc.; odour is covered under health, above);
- visibility;
- colour.

For pollutants for which there are numerous studies critical to the assessment, details of the individual study design should be presented in a tabular format with an overview presented in the text. Table 1 illustrates an ideal format; in practice the information available may not permit completion of every row or column. Presentation of the specifics of each study in this way provides a summary of the data, and shows the methods used to generate the data. Doses which had no significant effect ($p = 0.05$) should be included in the table, as this also facilitates derivation of the Reference Level. In those cases where there is a wide range of exposure times (hours, days, weeks, months), the data should be separated into short-term (acute), intermediate-term (sub-chronic), and long-term (chronic) subsets, and a table should be presented for each of them. To facilitate comparison between studies, it may be useful and necessary to provide some form of normalization of the exposure conditions in the studies used.

Table 1 Summary of scientific studies and results

Study no.	Study author(s), year, journal	Species, cultivar, race, life, stage, age	Concentration, exposure time, dose	Protocol (std. Protocol, other experimental details)	Endpoint measured	Significant ($p = 0.05$) effect, or no effect	Variability (SD, SE_x, etc.)
1							
2							
etc.							

Following a thorough evaluation of the compiled data, select the key study(ies) and critical effect(s). Identify the maximum dose at which no adverse effects were observed. For each endpoint, identify or calculate a No Observed Adverse Effect Level/No Observed Effect Level (NOAEL/NOEL) (section 11.2); in the absence of sufficient data for identifying the NOAEL/NOEL, identify or calculate a Lowest Observed Adverse Effect Level/Lowest Observed Effect Level) (LOAEL/LOEL). Providing the data on which the NOAELs/LOAELs are based are of sufficient quality to be used for risk characterization, consider effects with NOAELs/LOAELs within one order of magnitude. Where possible, derive dose–response relationships above the threshold values.

As is so often the case in risk assessment, it is not possible to provide an exhaustive list of considerations or a "checklist" approach to a qualitative understanding of the database. However, it may be constructive to consider the following in weighing the evidence:

- Are responses replicated in more than one species, by different investigators?
- Can the results be extrapolated from one population to another, across sex, strain, species, and pathway?
- Are the responses consistent across target organs, physiological endpoints, morphological attributes, etc?
- Is the toxicological data relevant to the expected exposure in terms of route, timing, frequency and duration of exposure (i.e., is there clear evidence of a dose–response)?
- Is there a plausible relationship between data on metabolism, postulated mechanism and the effect of concern?

It is important to critically evaluate apparent inconsistencies in the data to determine whether they represent significant biological differences, or whether they can be explained by differences in study design or other factors. Using this approach, the major strengths and weaknesses of the assessment can be conveyed so that they outline the availability of data and its current limitations. It is also important to provide information on the consensus — or lack thereof — within the scientific community. If the evidence is inconclusive, ambiguous, or equivocal in weight, the assessors should include their views on alternative approaches to the dose-response evaluation and qualitative factors to be considered in the risk characterization.

Provide qualitative conclusions about the likelihood that the substance may pose a hazard to human health or the environment, the nature and severity of the effects, and the conditions (route, dose levels, time, and duration) of exposure under which these effects occur. Finally, it is useful to provide a recommendation for future research, or a description of research in progress, which may provide clarification of uncertainties noted in the assessment.

11.2 QUANTITATIVE ASSESSMENT

The quantitative assessment in its simplest form is a numerical expression of the dose–response relationship. This step occurs when the assessor is satisfied that such numerical expression is supported by the underlying evidence and theory contained in the available data.

There are two general categories of endpoint in which this numerical expression occurs: threshold and non-threshold. Threshold

endpoints are classically defined as those which have a measurable level below which there is no discernible effect on an organism or population. As well, the theory of the mode of action of these toxicants would lead the assessor to conclude that adverse effects below this level would not be expected. Above this "threshold," there is a dose–response relationship which can be specifically quantified. Common expressions of this threshold concept are given by the NOEL, NOAEL, or the benchmark dose. Where a threshold is not actually observed in a study, there may be an expression of the LOEL where other evidence (other experiments, underlying mechanism, PBPK modelling) indicates the likelihood of a threshold's existence. LOELs (or LOAELs) may require the application of a specific uncertainty factor to bring them into the range of the NOEL.

Non-threshold endpoints are those in which the effect is proportional to concentration with no definitive loss of the relationship (i.e., no concentration at which the effect ceases to be observed) as exposure concentration decreases towards zero. The classic non-threshold endpoints are those such as occur with many carcinogens where the study cannot detect the presence of a threshold, and there is at least a theoretical basis for not assuming one (e.g., the substance of concern causes the effect by acting directly on genetic material). A common characteristic of this type of study is that the observations occur in a relatively high dose range and require extrapolation to an exposure level relevant to the general population of a species. This type of study is common in laboratory animal toxicology studies, and human occupational exposure studies. The important concept to be communicated in these studies is that of incremental risk (i.e., the expected increase in adverse outcomes with increases in exposure).

The steps in quantitative characterization begin with a characterization of the dose-response patterns for the effect(s), and description of the shapes and slopes of the dose-response curves for the various endpoints and receptors. Quantify the strength with which a particular receptor response is caused by the substance at various levels of exposure or dose.

In the absence of human data, use animal toxicology studies to derive estimates of effects, dose–effect and dose–response relationships. Data from animal studies should be converted to estimates of dose rate at the receptor of concern. If feasible, use pharmacokinetic models accounting for differences in deposition, clearance, lung retention patterns, metabolic rates and organ size, and chemical transformation to provide more relevant estimates of effect(s). For extrapolation from animal studies to humans (to the extent possible), the most relevant animal species should be used to generate dose-response data.

Generate dose-response curves by:

- extrapolation from doses in the experimental range to low dose levels of expected receptor exposure; and
- extrapolation, as required, from animals to humans, between laboratory species, from lab species to environmental biota, from high to low exposure situations, from short-term to chronic exposure, from single to multiple chemical exposure, from one chemical to another.

Use several models, if feasible, to extrapolate from high to low dose and across species. Available human data can be used

to validate and calibrate dose-response estimates derived from animal data.

Display the dose-response curve as incremental increases (or decreases) in response with dose (i.e., display graphically or in tabular form the probability of experiencing a particular health effect). Account for the statistical uncertainty in the results of the animal tests. Provide information on the variability in receptor responses for which the risks are to be estimated. Identify the dose which produces no detectable response above background, thus providing a NOEL. Alternatively, a useful approach developed recently has been to derive a so-called benchmark dose. This method involves fitting a model to the dose-response data that will allow calculation of confidence intervals on the dose–response relationship, interpolation between empirical observations, and extrapolation towards the origin below the lowest measured response. The benchmark dose is established by determining the upper bound estimate predicted by the model for a specified response level, such as 5% or 10%. The procedure will, in effect, generate a lower-bound confidence on the dose for that specific response level. The benchmark dose is therefore a defensible estimate of the minimal likely exposure required to produce a specified effect. There may be several benchmark doses, each specific for a particular effect.

Dose–response relationships should be determined for environmental effects as well. It is potentially more difficult for an environmental ecotoxic assessment because extrapolation of data from a few test species studied under lab conditions to the multitude of species in the natural environment is required. However, every attempt should be made to develop estimates of dose–effect and dose–response relationships for environmental receptors.

Identify and, to the extent possible, quantify the sources of variability and uncertainty. Variability should include major factors such as the variability in organism response (e.g., biological variation), and variability of exposure levels. Some sources of variability are a function of dose, such as latency period for cancer, and these should be noted where known. Uncertainty analysis should include such major factors as confidence in dose-response model extrapolations, including the sensitivity of the model to different assumptions, and the absence or effective control of bias and confounding in the results.

12 EXPOSURE ASSESSMENT

There is generally a need to evaluate all routes and media of exposure. For some air pollutants (e.g., ozone, particles) there are single routes of exposure and other media need not be investigated. For others (e.g., formaldehyde, PAHs), a multi-media approach is required since other media may contribute significantly to overall exposure. Risk characterization and final derivation of the Reference Level may require an in-depth exposure assessment in this latter case. Although data describing occupational exposure are not directly relevant to the environmental risk characterization process, they will be useful in deriving the dose–response relationship for humans and for understanding the mechanisms of toxic effects.

Based upon the ambient data review (section 1) and adding additional data to improve spatial and temporal resolution where available and appropriate, establish the exposure of Canadian receptors to the air contaminant. Evaluate the scientific literature on exposure, measured using the presence of a chemical as a surrogate for exposure, direct monitoring of exposure (personal sampling or media sampling), and biological samples to establish past exposure levels among populations. A search for data on exposure of some environmental receptors (e.g., microorganisms, invertebrates) should also be made and, if available, included in the assessment.

Even when the Reference Level derivation does not require it, exposure assessment provides essential information for the characterization of risk. Such an assessment can provide insight into the exact nature of exposure and help in the development of f

ocussed management strategies. It may, in elucidating exposure issues, also serve to drive some aspect of public debate. The depth and accuracy should be tailored to provide the degree of knowledge required to support the risk estimation.

12.1 STUDIES AND SURVEYS OF POINT-OF-CONTACT EXPOSURES

Retrieve and evaluate the available information on actual exposure measured through the spatial proximity of ambient monitors to receptors. Determine the quality (precision, accuracy, completeness, etc.) of the data, and whether it is statistically representative of exposures in Canada.

Test the following questions:

- Are the sources and pathways of exposure described? An estimation of the relative contribution to exposure via different media is required.
- Have any pathways been overlooked, and if so are they significant?
- Do the methods and data analysis support the conclusions?
- Were controls in place to ensure quality data collection?
- Has the possibility of additive pathways been considered?
- Is the exposed population representative of the general population?
- Does the study describe the data collection method, sampling statistics, analytical methods, and data analysis procedure(s)?

Provide descriptions of how the receptor(s) is exposed to the substance. In animal studies, the routes of exposure should be summarized (e.g., inhalation, ingestion, dermal absorption); for plant studies, the routes would include stomatal uptake, or uptake from the soil via the roots. Describe the magnitude (pollutant concentration), frequency (how often exposures occur — daily, weekly, seasonally) and duration (minutes, hours, days, lifetime) of exposure. Review available information on the relative contribution of various sources to receptor exposure. Explain if and how the exposure profiles are changing with time, and the sources which contribute to most of these changes. Determine if short-term exposure data are representative of a longer period.

Assess the spectrum of receptors exposed to the air contaminant, and identify the subset of these receptors likely to exhibit heightened sensitivity. Describe any activity patterns which may lead to increased or decreased exposure, and describe the receptor subset which is at highest risk resulting from these exposures, either through biological susceptibility (pre-existing conditions and physiologic assessment) or at the high-end of the exposure distribution. Identify any personal, cultural, regional, or socio-economic factors that can affect the duration and frequency of exposure for an individual or a population (i.e., the length of time one lives in a home or apartment, amount of time spent outdoors by persons of different ages, type of employment). Similarly, for environmental receptors, a description of the conditions under which enhanced sensitivity is likely to be present is required. This would include cultural practices in agriculture and forestry, and livestock husbandry, which may lead to enhanced sensitivity in the receptor. A description of the extent of heterogeneity

within receptor categories should be presented.

Describe models, whether they are deterministic, empirical, or statistical. State any validation of the model and statistical assumptions underlying it. Explain uncertainties, combining both sampling variability calculated from the data and non-sampling errors caused by the model, and parameter assumptions.

Gather and evaluate data on the concentrations of the substance, or metabolite, in receptor tissues. Assess the validity of the pharmacokinetic relationship of the measured biomarker. If available, present information on the absorption, distribution, and metabolism within, and elimination of the substance from the receptor.

12.2 EXPOSURE SIMULATIONS

Given quantitative data on the concentration and distribution of the substance, collect and evaluate the data on population characteristics of the receptor (demographic data, activity diary information, ecological range, distribution, etc.). Calculate the degree of exposure of humans and environmental receptors from various media.

Display integrated exposures for the population, population segment or receptor classes in histograms or curves. Identify the potential for high exposures. Highlight the uncertainties in the exposure estimates and the relative importance of the key assumptions and data. Provide estimates of the central tendency, upper and lower bounds on exposure, or the full population distribution.

For modelling human population exposure, obtain data on source emission rates, ventilation and infiltration, removal by adsorption onto surfaces, mixing, volume of space in which exposure occurs, activity patterns of individuals in each of the environments being modelled, demographic and census information, and information on the relationship between ambient and indoor concentrations. Produce probabilistic exposure estimates for representative populations, and describe the basis of the model and any validation status. Identify any research or data that is necessary to improve the modelled exposure estimates. Similar information should be gathered for environmental receptors. Additionally, land use data (agriculture, forestry, recreation, etc.) may be required to model receptor exposures.

13 RISK CHARACTERIZATION

Risk characterization is based on the best available information on probable exposure levels in authentic situations compared to levels believed to cause adverse effects. The analysis is based upon the state of knowledge regarding the nature and degree of hazard posed by the substance at known exposure levels. This process is an application of the “art” of risk assessment because inference and judgement must often substitute for verifiable knowledge. Given the sensitivity of this approach to the underlying assumptions and chain of logic, it can be reliable only when the underlying scientific theory and evidence is sound. Decisions should be heavily influenced by biologic factors and scientific judgement rather than over-stating or under-estimating the significance of mathematically modelled data. An accurate and unbiased discussion of the significance of the data is required, as is information on a variety of endpoints that provide insight into the full spectrum of responses in a number of receptors.

13.1 CHARACTERIZATION OF EFFECT(S)

13.1.1 Non-threshold Effects

Characterize the dose-response patterns for the effect(s) by discussing the slopes of the dose-response curves for the various endpoints and receptors. Current practice is to apply a non-threshold model to exposure situations in which there is no known or surmised threshold effect, such as carcinogens; in such cases the most conservative assumption is a linear extrapolation through the origin. The new EPA Carcinogen Assessment Guidelines (1992) are abandoning the linear fed multistage model as arbitrary and not defensible in favour of a model-free linear extrapolation.

Combine results of exposure assessment and the relevant dose-response relationship; describe what physiologic processes critically affect disposition kinetics, and how these processes vary with age, time and within the population. Determine whether increased susceptibility may occur as a result of concurrent exposures to other agents (interactive effects), concurrent disease, nutritional status and/or environmental conditions.

Identify sources of uncertainty, including statistical sampling issues concerning environmental data, dose-response models and their input parameters, and incomplete understanding of the biological cause and effect relationship with the air contaminant.

13.1.2 Threshold Effects

Thoroughly evaluate the data to select the key study(ies) and critical effect(s). Provide qualitative, weight-of-evidence conclusions about the likelihood that the substance may pose a hazard to human health or the environment, the nature and severity of the effects, and by what route(s) these effect(s) are seen to occur. Include information on the rationale behind the determination of the NOAEL/LOAEL, and any underlying assumptions.

Providing the data on which the NOAEL/LOAELs are based are of sufficient quality to be used for risk characterization, consider effects at the LOAEL when this occurs within one order of magnitude of the NOAEL, taking into account uncertainties; and discuss whether it is likely that alternative study designs or methods might produce a LOAEL that would encroach on the apparent NOAEL. If so, identify any need to presume a NOAEL lower than that empirically demonstrated.

For exposures for which a threshold is apparently lacking, a benchmark dose may be derived that is the lower confidence limit on a dose corresponding to an increase in the incidence of an effect at a particular level (e.g., the LED10 is the lower confidence limit on an effective dose that produces a 10% increase in response). The basis for the selection of this effective dose should be justified. Benchmark dose does not calculate such risks according to current proposals. This approach is not intended to replace the appropriate use of a NOAEL when an exposure demonstrates an apparent threshold.

13.2 RISK ESTIMATION

Develop estimates of the percentage of the receptor population, or the number of persons, above a specified level of risk, NOAEL, or other level of interest. Develop distributions of these estimates from which policy relevant proportions may be derived (e.g., upper 95% estimates of point values). This may be simulated by methods appropriate to the model and the availability of data from which to make the estimation (e.g., from an exposure model (pNEM [probabilistic national air quality standards exposure model]) or from a Monte Carlo simulation) but the method used must be justified.

Present the following items in a clear and concise manner:

- the ratio of the exposure level to the threshold dose which gives some indication of the likelihood of occurrence of the adverse effects associated with exposure to the substance;
- a point estimate of risk for receptors (individuals/species). This is an estimation of the number of individuals, species, strains or other subset of the population exposed to levels of concern;
- a point estimate or range of risk for a given population;
- probabilistic estimates of the population distribution of receptor effect cases over a specified time period. This is obtained either by summing the individual risks over all the individuals in the population; or by multiplying the slope factor obtained from a dose-response relationship, the arithmetic mean of the dose, and the size of the population; and
- estimates of the central tendency, upper and lower bounds.

Display all relevant information, including the nature and weight of evidence for each step of the process, the estimated uncertainty of the component parts, and the distribution of risk across various sectors of the population. Describe the upper end of the exposure distribution — the number of individuals in the most highly exposed group. Discuss the potential for exposure at still higher levels, and identify sensitive or susceptible populations. Ensure that the estimated high-

end exposure is within the expected distribution.

14 DERIVATION OF THE REFERENCE LEVEL(S)

The Reference Level is defined as a level above which there are demonstrated effects on human health and/or the environment. It provides a basis for establishing goals for long term air quality management. Following the evaluation and selection of the applicable scientific literature, integration of knowledge, and characterization of exposure and risk, the Reference Level(s) is determined. The Reference Level(s) provides the scientific basis for the establishment of a NAAQO, and is based on (a) the scientific evidence regarding the effects of the air contaminant on human health and the environment, and (b) the uncertainty (variability) associated with the compiled data set. No safety factor or other protective margin is incorporated into the Reference Level. Rather, it is intended to be the most appropriate summary benchmark for relating receptor effects to alternative proposals for objectives. Control technology, economic impacts associated with the level, and other management considerations are introduced in the next phase, which is the derivation of the Air Quality Objective.

14.1 ASSESSING THE SCIENCE

To be included in the determination of the Reference Level, a study must satisfy the criteria set forth in preceding sections, and:

- the receptor used should be present in Canada, should be related to a receptor present in Canada, or extrapolation of the measured endpoint to an effect on a Canadian receptor should be possible;
- the conditions under which the experiment was conducted occur, or are expected to occur, in Canada; and
- there is general acceptance or consensus in the scientific community regarding the quality of the results and conclusions.

This process requires expert judgement, and application of this process will vary among air contaminants for which NAAQOs are being derived. The rationale used to remove an individual scientific study from consideration at this stage must be clearly stated.

If there is little or no consistency in terms of treatment concentrations and duration of exposures among the scientific papers, some means of normalizing the terms of the exposure assessment is required before comparisons can be made. Calculation of a dose for each experimental treatment is the most fundamental means of normalization. There are, however, a number of other methods, such as incremental expression of increasing risk or "elasticity," which measure the rate of increase in a measured effect as dose increases from the mean, and a means of describing the change in the slope of the curve. The best form of expression is often dictated by the nature of the data. What needs to be kept in mind at this stage is that the reader must be able to easily compare the response in endpoints for the variety of conditions presented in the database.

Present a table which includes all selected studies (see sample table 1, described in section 11). As described above, normalize the data within each table in some way for comparative purposes. Graphical representations may also prove useful,

however, this will only become apparent upon analysis of the database.

Information presented in this manner represents the Weight of Evidence. In cases where more data are available and relevant to the Canadian situation, and no single study stands out, the entire database can be used for the calculation of the Reference Level. This Body of Evidence Approach will involve some recalculation of relevant studies. The Body of Evidence Approach uses some numerical representation of the entire database to develop an expression of the Reference Level.

For some air contaminants, there may be one study that is particularly relevant to Canada, there may be only one study in which environmentally relevant air contaminant concentrations were used, or one study stands out as being the most comprehensive or applicable in the Canadian context. In such cases, a Definitive Study Approach to the derivation of the Reference Level is warranted. In this approach, a single study is selected as best representing the dose–response relationship and the conditions expected in the Canadian environment. In both cases, present the Weight of Evidence in summary format to substantiate the method used to select the Reference Level.

Selection of the correct approach will rely on expert judgement. This judgement will be critically evaluated by both the peer reviewers and by the WGAQOG to ensure that the most appropriate option has been chosen.

14.1.1 Definitive Study Approach

A Definitive Study is one which meets the following criteria:

- it is the only useful study on the effects of the air contaminant within the Canadian context; and
- it has used air contaminant concentrations, exposure periods, and environmental conditions which are environmentally relevant.

Base the rationale for selecting this approach on sound reasoning and state it clearly.

14.1.2 Body of Evidence Approach

This approach is used when there are many quality studies upon which to base the dose–response relationship. Consider the data from all acceptable studies in the derivation of the Reference Level. This is a difficult process, and will rely heavily on expert judgement.

As in the Definitive Study approach, tabulate the data and, if necessary, following normalization, plot them separately into acute, sub-chronic, and chronic data sets as required. Clusters of data points (or other trends) apparent in this plot will facilitate the selection of the studies to be used in the derivation of the Reference Level.

Due to the number and diversity of studies considered in this approach, it is inevitable that there will be uncertainty associated with the derived dose and ambient concentration. The Reference Level derived in this manner will be the best estimation from a number of different tests. Due to variations among studies, application of an uncertainty factor may be required.

14.2 SCIENTIFIC UNCERTAINTY AND THE APPLICATION OF UNCERTAINTY FACTORS

Setting environmental quality criteria for the management of environmental and human health risks must inevitably confront substantial uncertainty. An essential aspect of confronting uncertainty is to distinguish between the contributions provided by variability as distinct from true uncertainty. These concepts are often confused because both can be represented by probability distributions and most environmental health issues will involve both.

Variability refers to the true differences or heterogeneity in some parameter that is measured or predicted. Variability can be better understood by collecting more evidence or data, but variability is fundamentally irreducible, except by narrowing the focus of interest to deal only with a subset or strata of the total range of values which may arise. For example, no amount of research can alter the true variability in human body mass or height within the population, but attention could be focussed on only one age group, thereby possibly reducing the range of variability that had to be considered.

True uncertainty refers to our lack of knowledge about the true value for a parameter or its specific mode of action. Uncertainty can only be resolved by collecting more evidence or performing research to better understand a parameter.

In the foregoing example, there is some uncertainly in our determination of any individual's body mass, but our measuring tools allow us to reduce this uncertainty to an insignificant contribution. However, most critical measures in environmental health experience substantial uncertainty and some measures (e.g., what is the health risk at levels of exposure to toxic substances substantially below practical epidemiological or toxicological detection) may be beyond determination.

In setting environmental criteria, variability must be distinguished from uncertainty because variability addresses issues such as which members of a population or an ecosystem may be at risk, while true uncertainty addresses how sure we can be

about the risk for any individual or species with specified characteristics. These are clearly not the same concerns and it is essential that environmental quality guidelines consider each of these aspects explicitly.

Each scientific finding is subject to uncertainty, only some of which may be expressed quantitatively. Empirical methods for deriving knowledge from scientific experiments or observational studies rely upon statistical inference to distinguish observed differences from differences caused by chance (random) variation. Statistical significance testing is performed to determine how small the chance is that chance variations could explain the observations. Typically scientists will regard observations which have less than a 5% likelihood ($p < 0.05$) of being explained by chance variation as being statistically significant. However, this does not mean that observers can have 95% confidence that the observed effect resulted from the hypothesized cause, because bias or confounding with other known or unknown factors may have produced the observation. Furthermore, a finding of statistical significance on any observation is entirely silent on the questions of practical significance of the observation (e.g., is the effect big enough to be of concern).

Use the above procedures to determine the dose upon which the Reference Level will be based. Convert the dose to an ambient air concentration for each period (acute, sub-chronic, chronic). To determine the Reference Level, estimate the level of uncertainty associated with this exposure and concentration and, if applicable, apply an uncertainty factor to the concentration to derive the Reference Level(s).

14.2.1 Uncertainty Within an Experiment

A measure of uncertainty within an experiment or trial is given by the amount of variability associated with the measured parameters. This variability may be expressed as the variance, standard deviation, standard error of the mean, coefficient of variation, confidence intervals, or other statistical terms. As variability increases, a greater difference between the control and treatment means is required, or greater numbers of measurements must be made, to establish statistical significance (i.e., a significant difference between the control and treatment).

This uncertainty (variability) is already accounted for in the statistical analysis and determination of significance. As a result, application of an uncertainty factor to each individual study is not required in the derivation of the Reference Level.

14.2.2 Uncertainty Among Experiments

Of greater importance in the process of deriving the Reference Level are the uncertainties involved in extrapolating from one receptor to another (e.g., from one plant species to another, or from rats to humans). Equally important are extrapolations from experiments conducted under controlled environmental conditions to Canadian conditions, comparisons of the data obtained from one set of treatments to another set where the same receptor was treated with the same compound but under different experimental conditions, or determinations of the level of the air contaminant which may cause damage made on the basis of a limited data set. It may, therefore, be necessary to apply an uncertainty factor to the value derived using the above procedures to determine the Reference Level. The magnitude of the uncertainty factor will reflect the variability in the compiled data set, with a larger uncertainty factor associated with a highly variable database. In this case, as in many others in this protocol, expert judgement will be required in deriving an uncertainty factor. Clearly state the rationale describing the need for an uncertainty factor, and the methods of deriving the uncertainty factor.

14.3 DERIVATION OF THE REFERENCE LEVEL(S)

The Reference Level is a level above which there are demonstrated effects on human health and/or the environment. The Reference Level is to be scientifically-based. Conceptually, it defines the boundary between the LOAEL and the NOAEL and is considered to be a benchmark for determining a level of exposure just below that most likely to result in a defined and identifiable but minimal effect. Both the ambient concentration and the associated uncertainty within the data set are to be considered in determining the Reference Level. Expert judgement is required in deciding whether to adopt the Definitive Study Approach or the Body of Evidence Approach in assessing the level of variability and uncertainty in the data set and in integrating these two components.

The Reference Level is to be distinguished from the so-called reference dose (RfD), which is a derived level at which continuous exposure would not be expected to result in any adverse effect in any but the most sensitive receptor. RfDs are derived by applying a safety factor (erroneously referred to in recent years as an uncertainty factor) to a LOAEL. In contrast, the Reference Level as used in this protocol has no safety factor applied to it and is related directly to the applicable LOAEL as described above. Conceptually, this means that continuous exposure to the Reference Level may be expected to result in adverse effects among receptors that are hypersensitive to the exposure, but the extent of susceptibility and the magnitude of the effect will vary depending on the exposure and the characteristics of the receptor. The Reference Level is intended to be a benchmark against which options for protection in the form of candidate levels for proposed National Ambient Air Quality Objectives may be compared.

15 PEER REVIEW

The Science Assessment Document will be reviewed by members of the WGAQOG. Additional internal reviews may be solicited as deemed necessary; these may include provincial departments of health and environment.

Following internal review, one or more external reviews may be commissioned. The external reviewers will be experts in the field(s) of study summarized in the Science Assessment Document, and will be knowledgeable about the scientific literature that supports the derived Reference Level.

These external reviews will be submitted to the WGAQOG, and the document will be revised as necessary.

Such reviews may take into account comparisons and compatibility with other relevant air quality objectives or standards, recommendations of international bodies, and the history of guidelines and objectives in Canada.

Responsibility for coordinating both the internal and external review processes rests with the WGAQOG.

16 GLOSSARY

Air Quality Objective:

The air quality management goal for the protection of the general public and the environment in Canada. It is a level based upon consideration of scientific, social, economic and technical factors.

Benchmark dose:

A defensible estimate of the minimum likely exposure that will produce a specified effect.

Biological monitoring:

Pollutant concentrations or metabolites are determined in exhaled air, saliva, blood, etc., as a measure of prior exposure.

Contact surface:

A part of the external surface of a target with homogeneous pollutant concentration.

Direct personal exposure measurement:

Participants carry monitors which measure the concentration encountered during daily activities, across different microenvironments.

Exposure:

Contact between a receptor and an environmental pollutant on a contact surface, at a certain concentration during a certain period of time.

Exposure profiles:

The collection of instantaneous exposures during a specified time period.

Exposure indicators:

Any measure or surrogate of true personal exposure to an environmental pollutant that is used in an epidemiological study to discriminate individuals on the basis of their exposure.

Ex vivo:

Tests conducted on excised tissues or organs.

Indirect personal exposure measurement:

Concentrations are determined in different media, and integrated personal exposure is calculated on the basis of dietary intake, time-activity data, etc., which are either estimated or determined.

Interim Air Quality Objective:

A tentative air quality objective based on an incomplete Reference Level data set and which, therefore, should be further researched and improved.

In vitro:

tests conducted in laboratory apparatus.

In vivo:

Tests conducted on organs or tissues within the living body.

LED:

The Lower confidence limit on an Effective Dose that produces a 10% increase in response to a chemical.

Non-threshold endpoints:

Dose-response endpoints which have a linear relationship.

pNEM:

U.S. EPA Probabilistic National Air Quality Standards Exposure Model is a personal air quality exposure model used to evaluate population exposure to pollutants.

Priority Substance List:

A formal CEPA process used to evaluate and recommend priority chemicals for management actions.

Reference dose:

A calculated or derived value (including a safety factor) which results in no effects with prolonged exposures.

Reference Level:

A level above which there are demonstrated effects on human health and/or the environment. It provides a scientific basis for establishing goals for air quality management.

Risk evaluation:

The process of interpreting risks, including levels of risk acceptable to individuals, groups or society as a whole.

Risk management:

The process whereby decisions are made about whether an assessed risk needs to be reduced to protect ecosystems and the means that should be used to achieve the desired reduction.

Science Assessment Document:

The complete compendium of the contents of sections 1 through 14 of this protocol which substantiate and recommend the Reference Level for a specific air quality pollutant.

Species:

A group of individuals or objects having certain distinguishing attributes in common; a genetically identifiable unit.

Strategic Options Process:

A process constituted under CEPA that develops management strategies for substances that have toxic exposure assessments.

17 BIBLIOGRAPHY AND REFERENCES

Canada (1986) The Queen's Privy Council. Citizens' Code of Regulatory Fairness, Regulatory Reform Strategy.

Covello, V.T. and M.W. Merkhofer (1993) Risk Assessment Methods: Approaches for Assessing Health and Environmental Risks. New York: Plenum Press.

de Koning, H.W., ed. (1989) Setting Environmental Standards: Guidelines for Decision Making. Geneva: World Health Organization.

Duan, N., A. Dobbs and W. Ott (1990) "Comprehensive definitions of exposure and dose to environmental pollution." Total Exposure Assessment Methodology. Pittsburgh: AWMA.

Federal, Provincial and Territorial Advisory Committee on Population Health (1994) Strategies for Population Health: Investing in the Health of Canadians. Ottawa: Minister of Supply and Services Canada, Cat. No. H39-316/1994E.

Lind, N.C., J.S. Nathwani and E. Siddall (1991) Managing Risks in the Public Interest. Waterloo, Ontario. Institute for Risk Research.

Pengelly, L.L., F.S. Silverman and S.M. Taylor (1994) A hierarchy of potential health effects from air pollution exposure. McMaster University and University of Toronto: Institute of Environment and Health.

U.S. Environmental Protection Agency (1992) Guidelines for Exposure Assessment. Federal Register, 57(104), 22888-22938 (FRL-4129-5).

U.S. National Research Council (1991) Human Exposure Assessment for Airborne Pollutants Advances and Opportunities. Washington, DC: National Academy Press.