Environmental Risk Assessment Guidance Manual

for agricultural and veterinary chemicals

Prepared by:
Chris Lee-Steere
Australian Environment Agency Pty Ltd

The goal of these guidance manuals is to provide a thorough and transparent discussion of the scope of what the Department of the Environment, Water, Heritage and the Arts considers when conducting risk assessments for industrial and agvet chemicals. These manuals provide a technical outline of the considerations that an assessor needs to take into account. To enhance stakeholder understanding of the environmental risk assessment process, a separate overview document is available.

The manuals provide general guidance material and are not intended to be exhaustive of every circumstance, however, it is expected that the manuals will be updated based on practical application and informed from stakeholder feedback. As such, these manuals rely on the professional judgement of the assessor and are not intended as a set of prescriptive or contestable mechanisms and processes.

Published by the Environment Protection and Heritage Council, February 2009
Hard copy: ISBN 978-1-921173-38-7
Web copy: ISBN 978-1-921173-39-4

ACKNOWLEDGEMENTS

Due to the comprehensive and readily accessible guidance material available internationally, this document has been written largely in an overview fashion, drawing on material from international jurisdictions, specifically, the US EPA and the European Union. Several source documents have been relied on and thanks are extended to the authors and organisations that have made this material readily available.

Some specific information resources and source documents in alphabetical order include:

- ECETOC Report on *Scientific Principles for Soil Hazard Assessment of Substances*
- European Union *Technical Guidance Document on Risk Assessment* (referred to throughout this manual as TGD)
- European Union *Guidance Document on Risk Assessment For Birds and Mammals*
- European Union *Guidance Document on Terrestrial Ecotoxicology*
- European Union *Guidance Document on Aquatic Ecotoxicology*
- Environment Canada's *Guidance Manual for the Categorisation of Organic and Inorganic Substances on Canada's Domestic Substances List - Determining Persistence, Bioaccumulation Potential and Inherent Toxicity to Non-Human Organisms*
- OECD's Workshop *Report on Statistical Analysis of Aquatic Toxicity Data*
- OECD's *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures*
- OECD's *Guidance Document on the Use of the Harmonised System for the Classification of Chemicals which are Hazardous for the Aquatic Environment*
- OECD's *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures*
- OECD's *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals*
- OECD's *Draft Guidance Document for the Statistical Analysis of Ecotoxicity Data.*
- OECD's *Report on Persistent, Bioaccumulative and Toxic Pesticides in OECD Member Countries*
- OECD's *Manual for Investigation of HPV Chemicals*
- OECD's *Report on Issues to be Addressed to Develop the Classification and Labelling for Terrestrial Environmental Hazards*
- US EPA's *Technical Overview of Ecological Risk Assessment*
- US EPA's *Initiative to Revise the Ecological Assessment Process for Pesticides.*

Many other source documents are cited throughout this manual and full details are found in the reference list.

TABLE OF CONTENTS

Chapter 1 – How to use this manual 7
1.1 Introduction 7
1.2 Terminology 8
1.3 General Discussion on Environmental Risk Assessment 8
1.4 International Best Practice 10
Chapter 2 – Planning the assessment 11
2.1 Introduction 11
2.2 The Planning Process 12
Chapter 3 – Data requirements 13
3.1 Introduction 13
3.2 Environmental Exposure Data Requirements 13
3.3 Environmental Effects Data Requirements 15
Chapter 4 – Data evaluation 18
4.1 Introduction 18
4.2 Determining the Quality of Data: (Reliability, relevance and adequacy) 18
4.2.1 Data reliability assessment 19
4.2.2 Determination of relevance and adequacy 21
4.2.3 Weight-of-evidence analysis 21
4.2.4 Monitoring data for existing substances 22
4.3 Use of Analogue Data 23
4.4 Use of QSARs 23
4.5 Expert Judgement 23
4.6 Data Reporting 23
Chapter 5 – Environmental exposure assessment 25
5.1 Introduction 25
5.2 Release Estimation 25
5.2.1 Volumes of use 25
5.2.2 Application and use pattern 25
5.2.3 Crop profiles 26
5.2.4 Formulation, handling and disposal 26
5.3 Environmental Chemistry and Fate 26
5.4 Exposure Calculations 28
5.4.1 Terrestrial organisms 28
5.4.2 Exposure estimates for non-target terrestrial plants 36
5.4.3 Groundwater 38
5.4.4 Exposure estimates for aquatic organisms 38
5.4.5 Exposure estimates for amphibians and reptiles 41
Chapter 6 – Environmental effects assessment 42
6.1 Introduction 42
6.1.1 Comment on evaluation of data 42
6.1.2 Comment on use of assessment factors 42
6.1.3 Comment on use of NOECs 43
6.1.4 Comment on emerging issues 43
6.1.5 Comment on formulations 44
6.2 Terrestrial ecotoxicity data 44
6.2.1 Birds 45
6.2.2 Mammals 46
6.2.3 Non-target invertebrates and soil micro-organisms 46
6.2.4 Non-target terrestrial plants 50
6.3 Aquatic Ecotoxicity Data 52

6.3.1 Introduction 52
6.3.2 Evaluation of data used for the assessment 53
6.3.3 Use of QSAR 53
6.3.4 Aquatic effects assessment 54
6.3.5 Fish 58
6.4 Assessment of secondary exposure effects 59
6.4.1 Introduction 59
6.4.2 Secondary exposure through the aquatic food chain 60
6.4.3 Secondary exposure through the terrestrial food chain 61
Chapter 7 – Assessment of Persistent, Bioaccumulative and Toxic (PBT) Substances 63
7.1 Introduction 63
7.1.1 International perspectives 63
7.1.2 Data gaps and other approaches 63
7.2 Overview 65
7.3 Data evaluation and available guidance 65
7.4 PBT Criteria 66
7.4.1 Persistence criteria 66
7.4.2 Bioaccumulation criteria 67
7.4.3 Toxicity criteria 68
Chapter 8 – Risk Characterisation 71
8.1 Introduction 71
8.2 Birds and Mammals 71
8.2.1 Acute (screening level) 71
8.2.2 Dietary (screening level) 71
8.2.3 Chronic (screening level) 71
8.2.4 Chronic (refined) 71
8.3 Bees 72
8.3.1 Standard risk assessment 72
8.3.2 Higher tier risk assessment 72
8.4 Other Arthropods 72
8.4.1 Current approach 72
8.4.2 EU approach 72
8.5 Earthworms and other soil macro arthropods 73
8.5.1 Standard risk assessment 73
8.5.2 Refined risk assessment for earthworms/soil macro invertebrates 74
8.6 Soil Micro-organisms 74
8.7 Non-Target Terrestrial Plants 74
8.7.1 International background 74
8.7.2 Current approach 75
8.8 Aquatic Organisms 75
8.8.1 Current approach 75
8.8.2 Sediment organism considerations 76
8.8.3 Higher tier risk assessment 76
8.9 Interpretation of Risk Quotients – Levels of Concern 76
8.10 Risk management options 78
8.10.1 Birds and mammals 78
8.10.2 Bees 78
8.10.3 Other arthropods 79
8.10.4 Soil micro-organisms 79
8.10.5 Non-target plants 79
8.10.6 Aquatic organisms 79
Chapter 9 – Veterinary medicines 80
9.1 Introduction 80
9.2 Protection Goals 80
9.3 General Description and Use of the VICH Guidance 81
9.4 Exposure of VMPs to the Environment 81

9.5 Risk Quotient (RQ) Approach....81
9.6 Test Guidelines....82
9.7 Metabolites....82
9.8 Special Consideration for Biodegradation Data....82
Chapter 10 – Probabilistic Risk Assessment....84
10.1 Introduction....84
10.2 OECD....84
10.3 European Union....85
10.4 United States....88
10.5 Australia....89
Glossary....90
REFERENCES....92
Appendix I – Relationship between various physico-chemical data....96
Appendix II – International Approaches to Aquatic Exposure Assessment....98
United States....98
European Union....102
Appendix III – Difficult to Test Substances....108
Appendix IV – Endocrine Disruption Potential....111
Appendix V – Bioaccumulation....114
Appendix VI – Higher Tier Aquatic Toxicity Testing....120

TABLES

Table 3-1: Data required for an environmental exposure assessment....14
Table 3-2: Data required for an environmental effects assessment....16
Table 4-1: Criteria for data reliability by type of assessment end-point....20
Table 6-1: Default BMF values for organic substances....60
Table 6-2: Assessment factors for extrapolation of mammalian and bird toxicity data....61
Table 7-1: Criteria for acute and chronic toxicity to aquatic species (algae, invertebrates, fish)....69
Table 8-1: Levels of concern for birds and mammals....76
Table 8-2: Levels of concern for terrestrial vertebrates, soil invertebrates, arthropods, and soil micro-organisms....77
Table 8-3: Levels of concern for terrestrial plants....77
Table 8-4: Levels of concern for aquatic organisms (including sediment organisms)....77
Table 8-5: Levels of concern for aquatic plants....77

FIGURES

Figure 1-1: Broad framework for conducting environmental risk assessments of chemicals....9
Figure 5-1: Soil accumulation in top 15 cm following 10 years of pesticide application at 150 g ac/ha....35

CHAPTER 1 – HOW TO USE THIS MANUAL

1.1 INTRODUCTION

The National framework for Chemicals Environmental Management (NChEM) is being developed under the guidance of the National Environment Protection and Heritage Council (EPHC). For further information on the environmental chemicals work being undertaken by EPHC please visit <http://www.ephc.gov.au/taxonomy/term/75>.

NChEM sets up a framework for environmentally sustainable chemical management, which consists of the following four linked action areas:

1) Environmental Risk Assessment – to make sure environmental risks from chemicals are identified and managed up-front and build in agency on-the-ground experience in the setting of management controls.

2) Environmental Controls – to bring national consistency to environmental regulation and management of chemicals and to ensure the right tools are available for the task.

3) Feedback of Information – to ensure chemical decisions are informed by on-the-ground experience and to improve the processes in place to gather, use and access this information.

4) Prioritising Action – to enable Environment Ministers to be proactive and strategically focused in identifying and addressing priority and emerging issues about chemicals in the environment.

Key area one focuses on improving environmental risk assessment of chemicals and this manual has been developed to improve transparency and understanding of enviornmental risk assessments.

In general, Australia seeks to keep pace with best practice in its assessment methodologies. This manual covers agricultural and veterinary (agvet) chemicals and is written based on current methodology in Australia and international best practice.

Its purpose is twofold. The first is to provide risk assessors with guidance on the environmental risk assessment of agvet chemicals. Secondly, it may provide other stakeholders with an illustration of the general process and considerations that risk assessors employ when assessing the potential risks that chemicals may pose to the environment. For a less technical illustration of the key areas that risk assessors consider, a separate overview document is available. The manual establishes a starting point for best practice assessment. It is intended that improved assessment tools and methods will be incorporated into the manual as they become available.

The manual provides information and guidance on various methods available to undertake the different areas of an environmental risk assessment, which, broadly speaking, comprises exposure, effects and risk characterisation. It outlines how the assessor should carry out an assessment of a new or existing agvet chemical according to best practice, including what information, methods and tools to use in assessing chemicals. Veterinary chemicals are discussed separately (refer Chapter 9) to agricultural pesticides, as agreed international methodology is available for the assessment of veterinary medicines. It is noted, however, that this information is provided as guidance rather than prescriptive metholodology as assessment needs to be tailored to fit the particular chemical being assessed.

The following chapters provide the assessor with the information that they need to carry out a risk assessment, including:

- General concepts on environmental risk assessment and the steps undertaken (this Chapter, Sections 1.2-1.3)
- Planning the assessment (Chapter 2)
- What data are required (Chapter 3)
- How data are evaluated for adequacy, suitability and reliability (Chapter 4)
- How environmental exposure is assessed (Chapter 5)
- How environmental effects are assessed (Chapter 6)
- How persistent, bioaccumulative and toxic chemicals are assessed (Chapter 7)
- How risk is characterised (Chapter 8)
- Assessing risks particular to veterinary medicines (Chapter 9)
- Probabilistic risk assessment (Chapter 10).

Such information provides a basis for a clearer understanding of the considerations that apply when assessing the potential risks that agvet chemicals pose to the environment.

1.2 TERMINOLOGY

The terminology used in assessments of industrial and agricultural chemicals has been developed through different streams over many years. In this manual and in the industrial chemicals manual the term predicted environmental concentration or PEC has been used for the concentration of chemical that is estimated to reach the environment. However, in agricultural chemical assessments this has historically been known as the estimated environmental concentration.

The OECD has provided a standard list of terms and abbreviations in their monograph guidance for plant protection products, and these terms will be used within this manual – refer to <http://www.oecd.org/dataoecd/46/43/1943970.pdf>. A full copy of the monograph guidance can be found at <http://www.oecd.org/dataoecd/45/60/1943906.pdf>

1.3 GENERAL DISCUSSION ON ENVIRONMENTAL RISK ASSESSMENT

In Australia, the Department of the Environment, Water, Heritage and the Arts (DEWHA) undertakes the environmental risk assessments of industrial chemicals for the National Industrial Chemical Notification and Assessment Scheme (NICNAS) and of agricultural and veterinary chemicals for the Australian Pesticides and Veterinary Medicines Authority (APVMA). There is a separate manual available that outlines the process whereby potential risks posed by industrial chemicals are assessed.

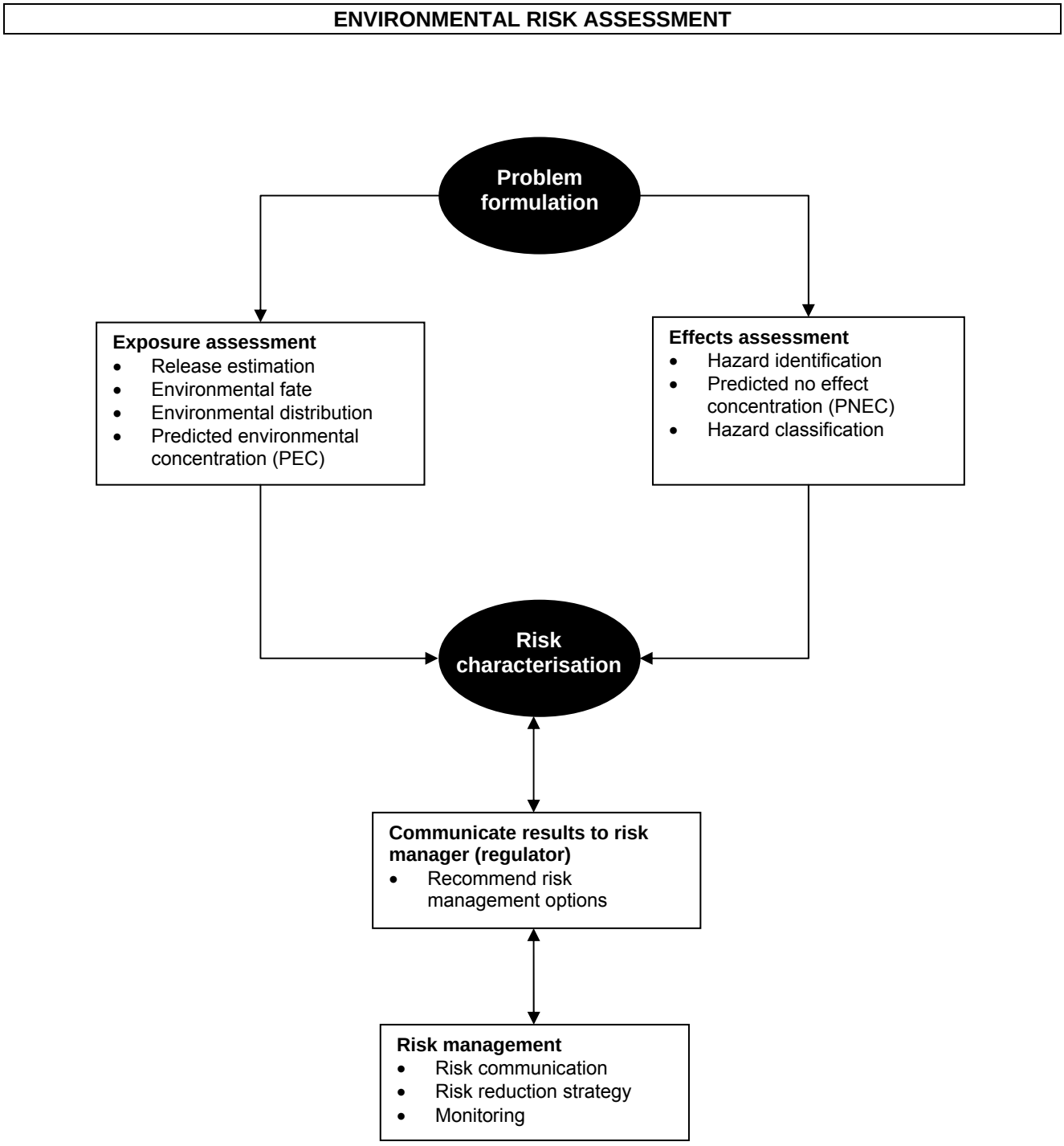
In 1983, the National Academy of Sciences[1] in the United States developed a four-step paradigm for risk assessment and risk management (US EPA, 2000) as follows:

- **Hazard identification**: examining toxicity data to determine effects of a chemical on health of humans or other organisms
- **Dose-response assessment**: extrapolating toxicity data from high dose studies to predict the likely effect of low doses of the chemical (also referred to as hazard characterisation)
- **Exposure assessment**: magnitude, frequency and duration of exposure to a chemical (for example, exposures from proposed or actual manufacture, use or disposal of a chemical)
- **Risk characterisation**: estimates potential for, and magnitude of, risk to an exposed individual or population.

The components of the risk assessment process are illustrated on the following page:

[1] NRC. 1983. *Risk Assessment in the Federal Government: Managing the Process.* National Research Council. National Academy Press, Washington DC.

Figure 1-1: Broad framework for conducting environmental risk assessments of chemicals[2]



[2] This framework is adapted from the US EPA framework for risk assessment which is representative of general methodology employed by Australia, Canada, USA and European agencies <http://www.epa.gov/oppefed1/ecorisk_ders/index.htm#framework>

This manual deals with the process to and including the risk characterisation. The further stages in the process, Risk Communication and Risk Management, are outside the current scope. However, some technical aspects of risk management are considered in this document where they can be used to refine the overall outcomes of the risk characterisation.

As explained in the United States Environmental Protection Agency's (EPA's) *Technical Overview of Ecological Risk Assessment* <http://www.epa.gov/oppefed1/ecorisk_ders/index.htm>, the assessment determines risk to plants and animals in the environment when exposed to a stressor such as a pesticide. In scientific terms, the risk assessment "evaluates the likelihood that adverse ecological effects may occur or are occurring as a result of exposure to one or more stressors". Undesirable events can include injury, death, or decrease in the mass or productivity of aquatic animals (eg fish and invertebrates), terrestrial animals (eg birds and wild mammals), plants, or other non-target organisms (eg insects), including endangered and threatened species.

- This assessment process combines all the information from the ecotoxicity tests (environmental hazard or effects), the exposure information, assumptions and uncertainties in a way that helps the risk assessor understand the relationships between the ecological effects and the stressors (pesticides). This process helps support decision-making.
- In determining whether a pesticide will harm the environment and wildlife, DEWHA conducts an environmental risk assessment for each pesticide active constituent including any major degradation products. Formulations are also assessed where relevant, including non-active constituents such as solvents, adjuvants and surfactants.

In Australia, environmental assessments of pesticides performed by DEWHA on behalf of the APVMA follow this broad approach. Individual chemicals are assessed on a case-by-case basis which means variations to the process are adopted where appropriate.

1.4 INTERNATIONAL BEST PRACTICE

The OECD Pesticide Assessment and Testing project works to harmonise the methods used by OECD countries to evaluate pesticide risks to health and the environment. This includes:

- developing test guidelines for the tests used to fulfil the pesticide registration data requirements
- harmonising exposure, hazard and risk assessment methods to interpret the test results and to assess a pesticide's risk.

Harmonisation means that the methods used amongst OECD governments are largely similar, though governments may retain some differences to account for national conditions and preferences.

In the areas of exposure, hazard and risk assessment, OECD governments work together to develop and harmonise methods to evaluate test results and other information, and to draw conclusions about hazards and risks. These become the basis for decisions on pesticide registration and risk reduction.

To this end, the OECD provides guidance documents to both governments <http://www.oecd.org/document/48/0,2340,en_2649_34383_2085104_1_1_1_1,00.html#monograph> and industry <http://www.oecd.org/document/48/0,2340,en_2649_34383_2085104_1_1_1_1,00.html#dossier> for chemical and biological pesticide registration. This guidance currently extends to formats for dossier presentation for industry, and assessment reports (monographs) from government agencies. While dossier presentation in Australia seldom follows that suggested in the OECD guidance, and assessment reports do not follow the current OECD monograph guidance, these are issues for the APVMA and outside the scope of this manual. It is noted, however, that both the APVMA and DEWHA are moving towards adoption of the OECD format.

Unfortunately, harmonised assessment methodologies are not available at OECD level. However, the risk based environmental assessment approach of Australia is very similar in structure to that of several OECD member countries including Canada, the USA and EU member states. These approaches, as described in this manual, can be considered as current international best practice.

For veterinary chemicals, a guidance document on environmental impact assessment of veterinary medicinal products has been developed through the VICH program (VICH, 2004). This internationally agreed approach provides a common basis for environmental impact assessments for veterinary medicinal products (VMPs) between the EU, Japan, US, Canada, Australia and New Zealand (refer Chapter 9).

CHAPTER 2 – PLANNING THE ASSESSMENT

2.1 INTRODUCTION

This step is often referred to as problem formulation and it is undertaken prior to the assessment itself. This process is applied in greater complexity with existing (review) chemicals than with the mandated assessment of new chemicals.

The management goal for new chemicals is clearly defined within the APVMA legislation. That is, that introduction of the new active constituent or product "would not be likely to have an unintended effect that is harmful to animals, plants or things or to the environment", or be harmful to the health and safety of human beings.

Existing chemicals are those chemicals that were already in use at the time that the APVMA was formed. These chemicals had been granted registration in Australia's states and territories under earlier arrangements, in some cases dating back as far as the 1950s.

Although existing chemicals should meet the same legislative requirements regarding likelihood of harm to people, plants, animals, things and the environment, they have not been subject to the same assessment and registration process as have new chemicals. Sometimes information becomes available which indicates that certain older chemicals may not meet contemporary standards for safety and therefore they must be reassessed using current regulatory standards.

This means that assessment of existing chemicals requires much greater problem formulation, and management goals may be more defined, particularly as the chemicals are already in use. Review chemicals are nominated and the outcomes of such reviews are very visible. The process for chemical review is available online at: <http://www.apvma.gov.au/chemrev/chemrev2.shtml>. It is simply summarised as follows:

- The APVMA becomes aware of new issues, for example, through new evidence being submitted, overseas regulatory action or through the Adverse Experience Reporting Program (AERP).
- These issues are considered and screened and the chemicals are chosen based on the strength of concerns relating to human health, occupational health and safety, the environment, trade and efficacy. This may involve seeking advice from external advisory Commonwealth agencies (for example, Department of the Environment, Water, Heritage and the Arts, Department of Health and Ageing) as well as jurisdictions, as appropriate.
- Once this initial screen has been completed, a review scoping document is prepared outlining the matters to be included in the chemical review.
- Companies that have registered products containing chemicals chosen for review are notified and required to submit any data they have about the chemicals relevant to the concerns.
- The APVMA seeks public submissions that address current uses of or problems with the products under review.
- Scientific data submitted is reviewed by the APVMA and external advisory Commonwealth agencies (for example, Department of the Environment, Water, Heritage and the Arts, Department of Health and Ageing).
- An evaluation of the use of the chemical including any potential risk mitigation strategies is released for government, public and industry comment.
- Comments relevant to the risk assessments are forwarded back to the external advisory agencies for evaluation, and further exploration of risk mitigation strategies may be pursued.
- A report on the review finding is published, which may also require that registrants submit altered labels with new directions for use.
- The APVMA takes the regulatory decision on the future use of the chemical under review.
- All participants in the review are notified of the APVMA's decision and regulatory actions are implemented. Outcomes of reviews are published in the APVMA Agricultural and Veterinary Chemicals Gazette, on the website and other media as needed.

This process dictates that the planning stage of the assessment needs to be very clear and transparent. To this end, the problem formulation steps outlined below play an important role.

2.2 THE PLANNING PROCESS

A detailed description of problem formulation may be found in the US EPA Guidelines for Ecological Risk Assessment (US EPA, 1998) or in their Technical Overview of Ecological Risk Assessment <http://www.epa.gov/oppefed1/ecorisk_ders/index.htm>.

Before the risk assessment is conducted, risk assessors (DEWHA) and risk managers (APVMA) engage in a planning dialogue. During the planning dialogue, which feeds into the problem formulation component of the risk assessment process, risk assessors and risk managers need to consider the following items:

Management goals: statements about the desired condition of ecological values of concern. For example, management goals may be to prevent toxic levels of contamination in water, sediments, and biota or to maintain a sustainable aquatic community. Management goals drive the risk assessment and often come from an enacted law or government policies.

Management options to achieve goals: risk management decisions that establish policy across the country for specific chemicals or sites. Management options may range from cancelling all uses of a particular pesticide to label restrictions for limiting the application of a pesticide in certain areas of the country. Management options define the scope, focus, and conduct of a risk assessment.

Scope and complexity of the risk assessment: risk assessments are constrained by the availability of valid data, expertise, time, and financial resources. The scope and complexity of the risk assessment is partly based on the uncertainty that can be tolerated in a decision supporting the risk assessment. Risk assessments that are conducted in support of legal mandates and likely to be challenged in court will require more resources and attention than those based on permitting a use of a pesticide in a small area.

After planning agreements are reached, the problem formulation phase, which is the foundation for the risk assessment, begins. In this process, risk hypotheses or assumptions are generated about why ecological effects may have occurred. Then the risk assessors and risk managers perform the following tasks:

Select assessment end-points: In a screening-level pesticide ecological risk assessment, typical assessment end-points are reduced survival, growth and reproductive impairments for individual animal species. For plants, the assessment end-points are typically concerned with maintenance and growth of non-target species. Although these assessment end-points are measured at the individual level, they indicate potential risk to populations.

Evaluate the nature of the problem: this step includes defining the nature of the stressor (pesticide) and characterising the pesticide use. In defining the nature of the stressor, risk assessors generally focus on the pesticide active constituent although in some cases they may consider pesticide formulations, inert ingredients, or degradates based on available data. Risk assessors use the pesticide product labelling to characterise the nature of the pesticide use in the field. Characterisation of pesticide use allows the risk assessors to focus the risk assessment on specific use patterns that are representative of a larger variety of use patterns. In this way, risk assessors can focus on use scenarios that reasonably represent the highest exposures.

Prepare a conceptual model: the conceptual model includes a set of risk hypotheses and a diagram that describe the predicted relationships among stressor, exposure, and assessment end-points. Typical conceptual models are flow diagrams that contain boxes and arrows illustrating these relationships. Developing a conceptual model allows the risk assessor to identify the available information regarding the pesticide, justify the model, identify data and information gaps, and rank model components in terms of uncertainty.

Develop an analysis plan: this is the final stage of problem formulation in which risk assessors develop a plan for analysing data and characterising risk. An analysis plan may summarise what has been done during problem formulation and target those hypotheses that are likely to contribute to the risk. It may also evaluate the risk hypotheses to determine how they will be assessed, develop the assessment design, identify data gaps and uncertainties, determine which measures will be used to evaluate the risk hypotheses (eg LC50, NOAEC, PECs), and ensure that the planned analyses will meet the risk managers' needs.

These plans may change and be refined as new developments and understandings arise during the assessment.

CHAPTER 3 – DATA REQUIREMENTS

3.1 INTRODUCTION

The following is a description of current data requirements for undertaking environmental risk assessments of agricultural and veterinary chemicals in Australia. These are reflected in the APVMA documentation to industry in their Manual of Requirements and Guidelines (MORAG), available at

<http://www.apvma.gov.au/industry/MORAG.shtml>

It is difficult to be too prescriptive in setting data requirements, given the wide range of applications in which agricultural chemicals may be used. The environmental data elements required for an application depend largely on the expected environmental exposure. All of the data requirements discussed in this chapter should be considered. Where data to address a specific requirement is not submitted, their omission must be justified with valid scientific argument, for example, by demonstrating that environmental exposure to this group of organisms will be minimal. Where these arguments are made and accepted by DEWHA, reasoning should be made clear in the assessment report.

The data elements that must be addressed for environmental risk assessment may be divided into:

1. Fate and behaviour in the environment (environmental exposure); and
2. Hazard: effects on non-target species (environmental hazard).

Environmental fate and behaviour

Environmental fate and behaviour data describe the degradation of active constituents, through abiotic and biotic mechanisms, and their mobility and likely transport and final destination in the environment. The data are used to help estimate the predicted environmental concentrations in different environmental compartments (vegetation, soils, sediment, water, air and animals), as appropriate, based on the proposed use pattern and physico-chemical properties of the chemical.

Environmental effects

Environmental effects data are obtained from tests on standard organisms, representing organisms that are likely to be exposed to the agricultural chemical product or to residues arising from its introduction into the environment. These data are used in conjunction with the anticipated environmental exposure and environmental fate data to determine the potential risk to non-target organisms, and the need for precautionary label statements or other risk management measures to minimise the potential for harm.

Non-target species

Tests for effects on non-target species include short term acute, subacute, reproduction, simulated field and full field studies. A hierarchical or tier system is used, under which the results from lower tier laboratory tests are used to determine the need for higher tier testing, such as full field studies, based on the potential for the chemical to cause harmful effects.

The APVMA requires that data submissions be made in accordance with the OECD common format for pesticide registrations. Under the agreed OECD format, the data should be presented in tiers. Submissions (dossiers) are required for active substances and formulated products. Details of the data elements required for environmental fate, behaviour and ecotoxicological studies, as well as supporting guidance, may be downloaded from the OECD website from the following URL:

<http://www.oecd.org/document/48/0,3343,en_2649_34365_2085104_1_1_1_1,00.html>

Individual data elements and the circumstances in which they are likely to be required are discussed in more detail in MORAG. This manual provides assessment guidance where data are supplied or required. For a better understanding of the likely data required for specific application types, MORAG should be consulted.

3.2 ENVIRONMENTAL EXPOSURE DATA REQUIREMENTS

The environmental exposure and fate of a chemical depends on its movement, transport and likely destination compartment. Australia is not a large manufacturer of agvet chemicals and so the primary focus of the exposure assessment for a pesticide often relates to end use. In some cases a more 'cradle-to-grave' approach may be taken to include the potential for release through manufacturing, formulation, and disposal of spent and unused product (for example, dipping solutions) and containers. Because many pesticides may be expected to be used up over time, disposal has not always formed a large component of assessments, however, disposal issues are now being considered more closely for some scenarios.

The following table provides an overview of the type of data that are important in assessing environmental exposure. Corresponding OECD data point numbers where available are provided. These are based on the OECD Guidance for Industry Data Submissions on Plant Protection Products and their Active Substances - Revision 2 May 2005; Appendix 6 - Format for the listing of test and study reports and other documentation; Part 4 - OECD, EU, US, Canadian, Japanese and Australian numbering systems for data and information on active substances. This document is available at: <http://www.oecd.org/dataoecd/43/25/34870442.pdf>.

In this regard, it is pointed out that the majority of data that will be received will be based on the active ingredient, not the end use product being assessed. In the following table, headings found in bold (including bold italics) show high levels where data are required to perform an assessment. The data end-points listed under these heading gives an idea of OECD data points that may be provided to address the overall requirement. Further, as explained in MORAG, not every data point below would be required or addressed for every application.

Table 3-1: Data required for an environmental exposure assessment

Data required	OECD Data Point Number
Extent of and potential for environmental exposure Amount of chemical to be used Manufacturing plant (active constituent) Formulating plant (product) Use and application Product disposal Accidental release	
Physical and Chemical Properties of the Active Constituent Melting and boiling points Relative density Vapour pressure and volatility Vapour pressure of the active constituent Henry's Law Constant Appearance Solubility in Water Solubility in Organic solvents at 15 to 25°C Partition coefficient n-octanol/water partition coefficient Effect of pH (4 to 10) on the n-octanol/water partition coefficient	**IIA 2** IIA 2.1 IIA 2.2 IIA 2.3 IIA 2.3.1 IIA 2.3.2 IIA 2.4 IIA 2.6 IIA 2.7 IIA 2.8 IIA 2.8.1 IIA 2.8.2
Abiotic degradation rate **Stability in water, hydrolysis rate, photochemical degradation, quantum yield and identity of breakdown products, dissociation constant.** **Estimated photochemical oxidated degradation** Hydrolysis rate of relevant metabolites Direct phototransformation of relevant metabolites	 **IIA 2.9** **IIA 2.10** IIA 7.5 IIA 7.6
Fate and behaviour in the Environment	**IIA 7**
Route of degradation in soil – laboratory studies	**IIA 7.1**
Aerobic degradation	IIA 7.1.1
Anaerobic degradation	IIA 7.1.2
Soil photolysis	IIA 7.1.3
Rate of degradation in soil – laboratory studies	**IIA 7.2**
Aerobic degradation of the active substance in soil (20°C and 10°C)	IIA 7.2.1 and IIA 7.2.2
Aerobic degradation, relevant metabolites, soil @ 20°C	IIA 7.2.3
Anaerobic degradation of the active substance in soil	IIA 7.2.4
Anaerobic degradation of relevant metabolites in soil	IIA 7.2.5

Data required	OECD Data Point Number
Field studies	**IIA 7.3**
Soil dissipation testing in representative soils	IIA 7.3.1
Soil residue testing	IIA 7.3.2
Soil accumulation testing on relevant soil	IIA 7.3.3
Mobility studies	**IIA 7.4**
Adsorption/desorption of the active substance	IIA 7.4.1
Adsorption/desorption of relevant metabolites	IIA 7.4.2
Column leaching studies with the active substance	IIA 7.4.3
Column leaching studies with relevant metabolites	IIA 7.4.4
Aged residue column leaching	IIA 7.4.5
Leaching (TLC)	IIA 7.4.6
Lysimeter studies	IIA 7.4.7
Field leaching studies	IIA 7.4.8
Volatility – laboratory study	IIA 7.4.9
Ready biodegradability of the active substance	**IIA 7.7**
Degradation in aquatic systems	**IIA 7.8**
Aerobic degradation in aquatic systems, metabolite identification	IIA 7.8.1
Anaerobic degradation in aquatic systems, metabolite identification	IIA 7.8.2
Water/sediment systems	IIA 7.8.3
Degradation in the saturated zone of the active substance and metabolites	**IIA 7.9**
Rate and route of degradation in air	**IIA 7.10**
Definition of the residue	**IIA 7.11**
Monitoring data concerning fate and behaviour (active and metabolites)	**IIA 7.12**
Other/special studies	**IIA 7.13**
Bioconcentration Potential in Fish[1]	**IIA 8.2.6**
Bioconcentration potential of the active substance in fish	IIA 8.2.6.1
Bioconcentration potential of relevant metabolites in fish	IIA 8.2.6.2
Aquatic bioavailability/biomagnification/depuration	IIA 8.2.6.3

1) Data related to bioconcentration in the OECD guidance document are noted under the general OECD data point number IIA 8.2 – Toxicity to Fish. However, these data relate to environmental fate and not environmental toxicity, so have been included in this table.

3.3 ENVIRONMENTAL EFFECTS DATA REQUIREMENTS

The following table provides an overview of the type of data that are important in assessing the potential for environmental effects. Corresponding OECD data point numbers where available are provided. These are based on the OECD Guidance for Industry Data Submissions on Plant Protection Products and their Active Substances -Revision 2 May 2005; Appendix 6 - Format for the listing of test and study reports and other documentation; Part 4 - OECD, EU, US, Canadian, Japanese and Australian numbering systems for data and information on active substances. This document is available at:
<http://www.oecd.org/dataoecd/43/25/34870442.pdf>.

Table 3-2: Data required for an environmental effects assessment

Data required	OECD Data Point Number
Ecotoxicological Studies	**IIA 8**
Avian toxicity	**IIA 8.1**
Acute oral toxicity (eg, quail, mallard duck)	IIA 8.1.1
Avian dietary toxicity (5-day) test, quail or mallard duck	IIA 8.1.2
Avian dietary toxicity (5-day) test in a second un-related species	IIA 8.1.3
Subchronic and reproductive toxicity to birds	IIA 8.1.4
Special field studies including palatability	
Fish Toxicity	**IIA 8.2**
Acute toxicity of the active substance to fish	IIA 8.2.1
Rainbow trout (*Oncorhynchus mykiss*)	IIA 8.2.1.1
Warm water fish species	IIA 8.2.1.2
Acute toxicity of relevant metabolites	IIA 8.2.1.3
Chronic toxicity to fish	IIA 8.2.2
Chronic toxicity (28 d exposure) to juvenile fish – growth/behaviour	IIA 8.2.3
Fish early life stage toxicity test	IIA 8.2.4
Fish life cycle test	IIA 8.2.5
Toxicity to aquatic species other than fish and aquatic species field testing	**IIA 8.3**
Acute toxicity to aquatic invertebrates	IIA 8.3.1
Acute toxicity for *Daphnia* (preferably *Daphnia magna*)	IIA 8.3.1.1
Acute toxicity for representative species of aquatic insects	IIA 8.3.1.2
Acute toxicity for representative species of aquatic crustaceans	IIA 8.3.1.3
Acute toxicity for representative species of aquatic gastropod molluscs	IIA 8.3.1.4
Chronic toxicity to aquatic invertebrates	IIA 8.3.2
Chronic toxicity to *Daphnia magna* (21-day)	IIA 8.3.2.1
Chronic toxicity for representative species of aquatic insects	IIA 8.3.2.2
Chronic toxicity for representative species of aquatic gastropod molluscs	IIA 8.3.2.3
Aquatic field testing (microcosms and mesocosms)	IIA 8.3.3
Effects on algal growth and growth rate (2 species)	**IIA 8.4**
Effects of sediment dwelling organisms	**IIA 8.5**
Acute test	IIA 8.5.1
Chronic test	IIA 8.5.2
Effrects on aquatic plants	**IIA 8.6**
Effects on bees	**IIA 8.7**
Acute oral toxicity	IIA 8.7.1
Acute contact toxicity	IIA 8.7.2
Toxicity of residues on foliage to honey bees	IIA 8.7.3
Bee brood feeding test	IIA 8.7.4
Effects on non-target terrestrial arthropods	**IIA 8.8**
Effects on non-target terrestrial arthropods/artificial substrates	IIA 8.8.1
Parasitoid	IIA 8.8.1.1
Predatory mites	IIA 8.8.1.2
Ground dwelling predatory species	IIA 8.8.1.3
Foliage dwelling predatory species	IIA 8.8.1.4
Effects on non-target terrestrial arthopods – extended laboratory, semi-field	IIA 8.8.2
Parasitoid	IIA 8.8.2.1
Predatory mites	IIA 8.8.2.2

Data required	OECD Data Point Number
Ground dwelling predatory species	IIA 8.8.2.3
Foliage dwelling predatory species	IIA 8.8.2.4
Other terrestrial invertebrates	IIA 8.8.2.5
Effects on earthworms (and other soil macroorganisms)	**IIA 8.9**
Acute toxicity to earthworms	IIA 8.9.1
Sublethal effects on earthworms and other soil macrooganisms (collembola)	IIA 8.9.2
Effects on soil microbial activity	**IIA 8.10**
Nitrogen transformation	IIA 8.10.1
Carbon mineralisation	IIA 8.10.2
Rates of recovery following treatment	IIA 8.10.3
Effects on marine and estuarie organisms	**IIA 8.11**
Marine or estuarine organisms acute toxicity	IIA 8.11.1
Marine/Estuarine fish – salinity challenge	IIA 8.11.2
Effects on terrestrial vascular plants (seed germination; vegetative vigour)	**IIA 8.12**
Effects on terrestrial vertebrates other than birds/wild mammal toxicity	IIA 8.13
Effects on other non-target organisms believed to be at risk	IIA 8.14
Summary of preliminary/range finding test data	IIA 8.14.1
Effects on biological methods for sewage treatment	IIA 8.15
Other/special studies	IIA 8.16
Laboratory	IIA 8.16.1
Field	IIA 8.16.2

CHAPTER 4 – DATA EVALUATION

4.1 INTRODUCTION

Not all data are created equal. Before data provided by notifiers can be used to assess the potential environmental effects of a new or existing chemical, the data must be checked for quality, including reliability, relevance and adequacy.

Any gaps in the data package should be identified and, if possible, filled. Because data can be expensive to generate, and recognising the push to limit animal testing, use of analogue data and modelling tools such as quantitative structure activity relationships (QSARs). This chapter guides the risk assessor in the process of evaluating data, filling some of the data gaps, and reporting of data.

Data quality directly influences how confident risk assessors can be in the results of a study and the conclusions they may draw from it. For example, specific concerns to consider for individual lines of evidence include whether the experimental design was appropriate for the questions posed in a particular study and whether data quality objectives were both clear and adhered to. Data may also be variable because of an inability to characterise the test substance properly, or different test guidelines compared with today's standards, or because certain information may have not been recorded that it has since been recognised as being important (Klimisch *et al*, 1997).

To this end, Australia prefers to receive test data that have been generated following established guidelines (eg OECD or US EPA), and performed following principles of good laboratory practice (GLP). Good laboratory practice defines a set of standards or principles for the planning, performance, monitoring, recording, reporting and archiving of a laboratory study — in short, the procedures necessary for the appropriate conduct of a physico-chemical or toxicological test. When these guidelines and practices are met then confidence in the quality of the study data are increased.

While it is most likely to be the case that test data for new pesticides and veterinary medicines follow standard guidelines and are performed according to GLP, this may not be the case for existing chemicals where data may be much older.

In these cases there is no harmonised international guidance available for evaluating data for agricultural chemicals. However, guidance is available for industrial chemicals. The principles are considered the same and the following chapter is based on the data evaluation chapter from the industrial chemicals assessment manual (Lee-Steere, 2005).

This chapter considers several aspects of data evaluation including use of analogue data and modelling tools such as QSARs in the event data are missing or questionable. The chapter is structured as follows:

- Data quality (Section 4.2)
- Use of analogue data (Section 4.3)
- Use of QSARs (Section 4.4)
- Expert judgement (Section 4.5)
- Reporting of data (Section (4.6).

4.2 DETERMINING THE QUALITY OF DATA: (RELIABILITY, RELEVANCE AND ADEQUACY)

This determination is particularly crucial in the case of information provided for an existing chemical. In such cases, information is most likely to have been generated prior to requirements of GLP and standardisation of testing methods. As well, data may be gathered from the scientific literature, relevant studies by current users of the chemical, previous registration packages, or from assessments undertaken by other regulators. For this reason, data for existing chemicals are much more likely to be variable in nature, requiring detailed data evaluation. Nevertheless, consideration of all available existing information for an existing chemical is important because, if it is judged to be of sufficient quality, there is no need for additional testing for that end-point, resulting in savings in resources such as time, costs and laboratory animals.

The process of determining the quality of existing data takes into consideration three aspects - adequacy, reliability and relevance of the available information to describe a given assessment end-point.

These terms were defined by Klimisch *et al*, 1997[3] along the following lines:

- **Reliability:** evaluating the inherent quality of a test report or publication relating to preferably standardised methodology and the way the experimental procedure and results are described to give evidence of the clarity and plausibility of the findings.
- **Relevance:** covering the extent to which data and tests are appropriate for a particular hazard identification or risk characterisation.
- **Adequacy**: defining the usefulness of data for hazard/risk assessment purposes. When there is more than one study for each end-point, the greatest weight is attached to the study that is the most reliable and relevant.

These terms and concepts are discussed further in the following OECD, 2004 (a) guidance on determining the quality of data (refer sub-sections 4.1.2 and 4.1.3). While it is stated in the OECD guidance that it is not intended to present all possible approaches which can be used to assess data quality, it does present two tools. One is already used by industry and another is commonly used by governments, which is a more criteria driven approach for compiling and assessing the completeness of data. In addition to this, guidance on determining the adequacy of data is available from the EU (EC, 1996 and 2003 (a)).

Each study essentially will require a case-by-case consideration and for these reasons a quick look at the reliability of the studies may save time later when relevance and adequacy are considered.

4.2.1 DATA RELIABILITY ASSESSMENT

The reliability of the data is a key initial consideration because without knowledge of how the study has been conducted all other considerations may be irrelevant. Screening for reliability can be done relatively quickly to filter out unreliable studies and enable the risk assessor to focus further resources on those studies considered most reliable.

The following approach for determining reliability is used by DEWHA:

1	**Fully reliable:**	GLP compliant and fully compliant with the Test Guideline specified.
2	**Reliable with restrictions:**	GLP compliant but not fully compliant with the Test Guideline specified, but nevertheless judged to provide a reliable basis for regulatory decision-making. An asterisk is to be added to identify studies that are not standard that are judged to be reliable for the purpose conducted (*e.g.* mechanistic studies)
3	**Not reliable:**	Not GLP compliant and/or not compliant with the Test Guideline specified, and judged to not provide a reliable basis for regulatory decision-making.
4	**Not assignable:**	Insufficient information provided to allow the reliability of the test or study report to be assessed (*e.g.* published literature).

It should be noted, these ratings are derived from the OECD. Australia does not have mandatory GLP and consequently some allowances need to be made in addressing the validity of a study. For example, non-GLP studies can not be considered unreliable on these grounds alone. Therefore, a degree of expert judgement has been used in applying the validity rankings associated with studies assessed.

The following guidance outlines a further two approaches, one developed by Klimisch *et al*,1997, and one developed by the US EPA HPV Challenge Program, which may be used as an initial or first screen of studies. Both are compatible and may be used either alone or together by assessors considering data quality. A description of both methods is currently provided in Chapter 3 of the OECD *Manual for Investigation of HPV Chemicals* – <http://www.oecd.org/dataoecd/60/46/1947501.pdf>

The approach by Klimisch *et al*, 1997 was developed as a scoring system for reliability, particularly for ecotoxicology and health studies (however, it may be extended to physico-chemical and environmental fate and pathway studies), as follows:

[3] Klimisch, HJ, Andreae, E and Tillmann, U 1997. *A systematic approach for evaluating the quality of experimental and ecotoxicological data*. Reg.Tox. and Pharm. 25:1-5

1 = reliable without restrictions: "studies or data...generated according to generally valid and/or internationally accepted testing guidelines (preferably performed according to GLP) or in which the test parameters documented are based on a specific (national) testing guideline...or in which all parameters described are closely related/comparable to a guideline method."

2 = reliable with restrictions: "studies or data...(mostly not performed according to GLP), in which the test parameters documented do not totally comply with the specific testing guideline, but are sufficient to accept the data or in which investigations are described which cannot be subsumed under a testing guideline, but which are nevertheless well documented and scientifically acceptable".

3 = not reliable: "studies or data...in which there were interferences between the measuring system and the test substance or in which organisms/test systems were used which are not relevant in relation to the exposure (eg unphysiologic pathways of application) or which were carried out or generated according to a method which is not acceptable, the documentation of which is not sufficient for assessment and which is not convincing for an expert judgement.".

4 = not assignable: "studies or data…which do not give sufficient experimental details and which are only listed in short abstracts or secondary literature (books, reviews, modeled results etc)." An exception may be in the case of peer-reviewed books such as the Merck Index and the CRC Handbook of Chemistry and Physics, to which the OECD Manual for Investigation of HPV Chemicals has allocated a reliability of code of 2. The assessor typically reads the data report, compares it to the relevant standard method and decides which code to allocate.

After assigning the relevant Klimisch code to each study, those with the lowest scores would be the most reliable. The use of Klimisch codes provides a useful tool for organising the studies for further review. For example, they enable the assessor to focus on the most highly reliable study first in order to allow time to later consider relevance and adequacy.

The second approach was developed in 1998 as part of the US EPA HPV Challenge Program, and provides more information than the Klimisch system by describing the key reliability criteria for each type of data (refer Table 4.1 for examples). These criteria address the overall scientific integrity and validity of the information in a study, ie reliability. This approach is consistent with the Klimisch approach - any study that does not meet the criteria in Table 1 would also not be assignable under the Klimisch system. Such studies may, however, be considered later as supplementary information to the overall assessment of a particular end-point, particularly if there is no single key study - with appropriate care due to potential limitations.

Table 4-1: Criteria for data reliability by type of assessment end-point

Criteria	Required for following information		
	P/Chem	Env. Fate	Ecotox
Test substance identification (Adequate description of test substance, including chemical purity and identification /quantification of impurities to the extent available).	X	X	X
Temperature	X[1]	X	X
Full reference/citation	X	X	X
Controls[2]		X	X
Statistics With some exceptions (eg the *Salmonella*/Ames assays).			X
Species, strain, number, gender and age of organisms			X
Dose/concentration levels		X	X
Route/type of exposure[3]			X
Duration of exposure		X	X

1. For vapour pressure, octanol/water partition coefficient and water solubility values
2. All studies must have negative controls and some studies (eg biodegradation) must also have positive controls. If a vehicle is used in the administration of the test agent, vehicle controls should be established and reported. Exceptions may be allowed for acute mammalian toxicity studies.
3. The route/type of exposure (eg oral inhalation etc for mammalian studies) or test system (static, flow through etc for ecotoxicity) must be reported.

Each study is evaluated against these criteria, allowing the assessor to set aside studies which fail to meet the essential criteria for reliability.

4.2.2 DETERMINATION OF RELEVANCE AND ADEQUACY

The next step is to determine whether the data are relevant, and whether they are adequate for fulfilling the needs in a hazard or risk assessment. The studies that have passed the initial screen for reliability should be considered. If crucial data are not reported then the assessor may be unable to determine whether the study can be used.

The use of sound scientific judgement is the most important principle in considering relevance and adequacy, because such a determination is so case-specific. For this reason there are no ranking criteria that can be listed as guidance. Nevertheless the following paragraph describes some of the considerations that assessors may apply.

Relevance is easy to establish in extreme cases. For example, data on appropriate Australian species in Australian conditions at realistic exposure levels of the chemical of interest are the most relevant of all. A more likely example of the consideration of relevance would be a situation where aquatic toxicity data have been generated on cold water fish that do not exist in Australia and whose preferred environmental conditions only exist in very few areas in Australia. When considering the potential environmental effects of a chemical under Australian conditions, such data would not be as relevant as data generated on warm water fish that may or may not exist in Australia but that fill a similar niche to Australian species and inhabit environmental conditions that are more common in Australia.

In some cases the type of substance under investigation will result in the recommended test for a particular end-point being difficult or inappropriate to carry out, eg, chemicals that are unstable in abiotic or biotic systems, chemicals with known explosive/flammable properties or volatile substances. In such cases the relevance of the study may be questionable.

Determination of adequacy depends on considerations such as the results found, the precision of the end points, whether studies differ in their results for the same test, the statistical power of the test, and how relevant the data are. Weight of evidence analysis (refer 4.2.3) also plays a role in the determination of whether the data package as a whole is adequate.

4.2.3 WEIGHT-OF-EVIDENCE ANALYSIS

The use of tools for identifying reliable data and expert judgement for determining relevancy and adequacy helps to ensure that high quality data are used. They do not, however, remove the need for a weight-of-evidence analysis approach during the assessment of these data.

Similarly, the assignment of Klimisch codes for data reliability does not necessarily mean that any extra weight should be given to these studies in the overall assessment, as there may be information from other studies on other end-points that have an influence. The assessment report should be explicit on the criteria, which have been applied to assess quality, rather than simply referencing a score.

Because of the nature of existing data, it is reasonable to expect that there will be some cases (for a given end-point) in which several studies - some of which may not have passed the initial screen - may be collectively used to fill the end-point, thereby avoiding additional testing. In other words, it may be possible to pool several studies, one or more of which may be inadequate in some way, to satisfy a specific end-point. For example, there may be several acute fish toxicity studies available on a particular chemical, none of which would be acceptable by itself due to some deficiency (ie low number of test animals/dose group, only one dose group in addition to control group, change in dose amount or frequency during the course of the study, etc.). However, if the different studies show similar effects and/or mortality at approximately the same dose and time, then collectively they could satisfy the toxicity data requirement.

It needs to be recognised that, for some substances, it may not always be possible to create a confident weight of evidence. For example, these substances may not have reliable experimental data, or they may be model difficult such that QSAR estimates are unable to be generated with confidence. In these cases, expert judgement must be used to determine the end-point.

4.2.4 MONITORING DATA FOR EXISTING SUBSTANCES

Monitoring data are available for air, water, sediment, biota and/or soil, for a number of existing substances. The EU Technical Guidance Document on Risk Assessment (TGD) provides useful guidance on selection of these data for use in undertaking an exposure assessment of existing chemicals, and the following discussion is paraphrased from this reference.

Monitoring data should be carefully evaluated for their adequacy and representativeness. They are used together with calculated environmental concentrations in the interpretation of exposure data. A stepwise evaluation procedure is recommended as follows:

- reliable and representative data should be selected by evaluation of the sampling and analytical methods employed and the geographic and time scales of the measurement campaigns
- the data should be assigned to local or regional scenarios by taking into account the sources of exposure and environmental fate of the substance
- the monitoring data should be compared to the corresponding calculated predicted environmental concentration (PEC). For naturally occurring substances, background concentrations have to be taken into account. For risk characterisation, a representative PEC should be decided upon based on measured data and a calculated PEC.

Experience in Australia has shown that Australian monitoring data are seldom available, so often international monitoring data are relied upon. The relevance of this to an Australian assessment is questionable, and often the best uses of these data are to help determine if the modelled environmental concentrations are realistic.

If monitoring data within Australia are available for a substance, their adequacy should be appropriately assessed and their use would be preferred to international data if of appropriate quality. *Measured concentrations that are not representative as indicated by an adequate sampling program, or are of insufficient quality, should not be used in the exposure assessment.*

The limit of quantitation (LOQ) of the analytical method should be appropriate for the risk assessment and the comparability of the measured data should be carefully evaluated. For example, concentrations in water may either reflect total concentrations or dissolved concentrations according to sampling and preparation procedures.

When a substance is used in materials (eg polymers) it may be released to the environment enclosed within some matrix of the material formed. In such cases it would be useful to know if the analytical method used is able to detect also the fraction of substance that is associated with these particles as it would affect availability of the chemical to the environment and its fate.

In selecting representative data for the environmental compartment of concern, there are two distinct aspects to consider:

- the level of confidence in the result (ie number of samples, how far apart and how frequently they were taken)
- whether the sampling site(s) represent a local or regional scenario.

It has to be ascertained if the data are results of ad hoc examinations or if the substance was detected at the same site over a certain period of time. *Measured concentrations caused by an accidental spillage or malfunction should not be considered in the exposure assessment.*

If there is no spatial proximity between the sampling site and point sources of emission, the data represent a regional concentration that needs to be added to the calculated local (PEC). If the measured concentrations reflect the releases into the environment through point sources, they are of a PEClocal type. In a PEClocal based on measured concentrations, the regional concentration can generally be assumed to be already included (although this may vary in some cases, such as where effluents are measured and then assumed to dilute).

Measured concentrations in biota may be available as samples of living organisms and may be used for environmental monitoring. They can provide a number of advantages compared to conventional water and sediment sampling especially with respect to sampling at large distances from an emission source or on a regional scale. Further, they can provide a PECbiota and consequently an estimation of the body burden to be considered in the food chain.

For a fuller discussion on the points raised in this section, Section 2.2 of the TGD can be consulted (EC, 2003 (a)).

4.3 USE OF ANALOGUE DATA

Argument may be made by applicants that some data do not need to be submitted in a specific situation. In some cases, data on a chemically similar substance (an analogue) may be submitted. Experience shows this is not a common approach for agricultural chemicals.

It is appropriate to investigate the use of analogues or surrogates to assist in providing supplemental data to reduce possible testing needs. In some situations data from another chemical can be used, such as:

- isomers which have similar structure activity profiles
- closely related homologues
- relevant precursors and breakdown products, along with information on metabolism and degradation.

The data for the related compound should be included in the assessment report for the chemical, clearly stating the identity (chemical name and CAS No.) of the related compound (test substance). When data for an analogue chemical are used to fill one or more end-points, the data for the analogue's other end-points must be compared and discussed in relation to the main chemical. This will shed light on the similarities and differences in the properties of the main chemical and its analogue (OECD 2004 (a)).

4.4 USE OF QSARS

Where experimental or analogue data are not available, values may be predicted using a suitable quantitative structure activity relationship (QSAR). In regards to predicting toxicity, QSARs can be applied to chemicals with a common mode of toxic action, such as narcosis where the mechanism is dependent on a chemical's hydrophobicity (eg, log Kow). However, agricultural pesticides or veterinary medicines are often biologically active with more specific modes of action, making the use of a QSAR approach less useful. Again, experience shows this is not a common approach with agricultural chemicals and consequently is not dealt with in detail in this manual. For a more detailed discussion on use of QSARs, the "Environmental Risk Assessment Guidance Manual for Industrial Chemicals" can be consulted.

When applying QSARs it should be taken into account that a QSAR is an estimation method and that therefore there is a certain probability that the estimate is poor, even for well evaluated models.

4.5 EXPERT JUDGEMENT

When no experimental data are available for a substance and predictions are not possible using QSARs (eg model difficult substances), expert judgement should be used. Environment Canada provides guidance for applying expert judgement (using rules of thumb) in their guidance document for categorising existing substances on their Domestic Substances List (DSL) (Environment Canada, 2003). Assessors should be aware of this manual and use it for guidance in the application of expert judgement where required during an assessment. The document can be obtained from Environment Canada by email at ESB.DSE@ec.gc.ca

4.6 DATA REPORTING

Reporting of data in the assessment report for new and existing chemicals should be transparent with a clear description of assumptions, choice of models, or expert judgements made in undertaking the assessment. When reporting test data, sufficient information should be included to give readers as clear an indication as possible of the test conditions, observations and results of the test. Where tests follow standardised and internationally recognised guidelines, the onus of reporting may be reduced regarding test systems and methodology. However, within guidelines, different options may exist for testing. Consequently, reporting of tests should include information as follows:

Test Substance: This refers to the identity of the chemical. Where possible the purity, percentages of known impurities, and details of any vehicle used in testing should be given. This is particularly important for existing chemicals where older test methods may be used and data are being evaluated from many sources. If the chemical used in the specific test was different from the commercial product (purity, additives, different solvent carrier, etc.), then those differences need to be noted. This notation should be included together with the chemical name and CAS number.

Method: If the study was done according to OECD test guidelines or other widely recognised standard test methods/guidelines then this should be identified. The year of publication of the guideline should be reported as well. In these instances only the name of the guideline needs to be reported, as a full description of the method is not needed. The same considerations apply for studies run under standard guidelines that have since

been superseded. When a non-standard method has been used, details of the method, equivalent to those in an OECD test guideline, should be provided. If such information is not available, then this fact should be noted.

When the test method allows the use of alternatives for certain test parameters (eg species), the study authors should have indicated the chosen alternatives. In the case of aquatic toxicity tests, it is important to indicate whether nominal or measured concentrations were used. If there have been deviations from the test guideline, then those deviations that will significantly impact either the study reliability or the interpretation of the data need to be individually listed. In cases where a single study addresses several end-points, the study may be reported for each with the results and conclusions sections differing depending on the end-point. However, the method and reference section would be the same in each case.

Test Conditions: Any relevant information on test conditions in a broader sense can be reported, ie test system including test conditions, testing procedure, temperature, pH, test system, etc.

Information on preparation of test solutions is very important, particularly with more insoluble substances. Use of solvents should be fully explained, or in cases where the water accommodated fraction is used, their preparation should be fully described and these should be properly identified in the results.

Results: At a minimum, qualitative descriptions of elements where dose-related observations were seen should be described and a NOEC and LOEC stated (where relevant) for critical effects together with the rationale for selection of these values (eg sub-lethal effects, mortality, etc.). In addition, if a study includes effects that were not considered to be biological or statistically significant, an explanation should be given. Expressing results by phrases such as insoluble in water is discouraged. A limit test should be performed under such circumstances so that a positive expression, such as <0.1 mg/L (analytical limit), can be entered. Calculated values must be identified and the calculation method should be cited.

Conclusions: The conclusions of the author of the study can be noted, together with any comments of the assessor. These should be clearly separated from each other by indicating the origin of the comments.

Reliability: This section can be used to denote the adequacy of data, at the discretion of the assessor, and is particularly important for data relating to existing chemicals. Data reliability codes can be used, as described in Section 4.1 above. The rationale for the reliability code should be described clearly, as should the process by which the reliability decision was made.

References: The name of the performing laboratory should be provided (or author/date if summarising a publication). The full reference should be given in the references section of the report.

CHAPTER 5 – ENVIRONMENTAL EXPOSURE ASSESSMENT

5.1 INTRODUCTION

In order to assess the potential environmental risks posed by a pesticide, it is crucial to identify how much of it is likely to reach the environment, where it will end up and how long it will stay there. This chapter provides guidance in assessing environmental exposure.

Environmental exposure assessments characterise either the extent to which organisms may be exposed to a chemical, or the concentration of a chemical in various environmental compartments, which may then have the potential to affect organisms. Exposure data are then compared to effects data in order to characterise risk. Effects data and how to assess them are discussed in Chapter 6.

There are three main steps to exposure assessment. These steps are:

1. **Release estimation** (eg how much will be used and where).
2. **Environmental chemistry and fate (**distribution) (eg will the pesticide degrade, break down, drift, bioaccumulate, etc. and where it will end up).
3. **Exposure calculations** (eg the amount of chemical predicted to be present in various environmental compartments).

5.2 RELEASE ESTIMATION

There are several factors that relate to how much chemical is available for release (volume) and ways a chemical or product could be released (application and use pattern) into the environment. Such information is used in determining the fate and expected environmental concentrations of the chemical in the various environmental compartments of the environment and includes assessment of:

- volumes of use
- application and use pattern (including rates and methods of application, number of applications per season and times of application)
- crop profiles
- formulation, handling and disposal.

Although these factors appear to be predominantly related to agricultural use, they are also relevant for home garden uses, as well as antifoulants.

5.2.1 VOLUMES OF USE

An estimation of the volumes of use of the new active is required. This information, together with the application and use pattern will allow a determination of the extent of environmental exposure.

5.2.2 APPLICATION AND USE PATTERN

A statement relating to the use of the product, the crops or resources it will be applied to, and the pests it aims to control must be provided. The application method should be described including route (ground, air), rate in g ac/ha; formulation type; spray intervals; number of sprays per season and any information available on tank mixing (eg mixed with water and applied in a minimum of 50 L/ha for ground application and 20 L/ha for aerial application). These details are also required for pesticides used around the home or for water treatment. Where applications occur by aerial or ground boom spraying, details on the proposed spray quality and droplet size must be provided.

Formulation types: For definitions and abbreviations of different formulation types and components, assessors should refer to the relevant section of MORAG (Vol 3, Part 2, Appendix B).

If the draft label contains critical comments relating to the application and use pattern, these details should be included in this Section. For example,

"Critical comments provided on the label advise that application occur at or just prior to the anticipated time of hatching of insect eggs as indicated by egg levels. When larvae are large (>5 mm), the addition of <product> is stated to possibly improve control".

5.2.3 CROP PROFILES

A crop profile should be included for each cropping situation outlined on the draft label. This profile should include information on geographic regions where such crops occur in Australia, likely soil and meteorological conditions and crop areas in hectares. This type of information contributes to the exposure assessment because it allows the assessor to refine how and where the pesticide may be released to the environment and how long it may persist. For example, depending on the particular pesticide, there may be a higher risk of movement of a water soluble pesticide to groundwater in regions where soils are sandy or permeable. If the crop is only grown in tropical regions, then this may indicate that the pesticide will be applied in warm and moist conditions, which may increase or change the rate at which the pesticide breaks down in the environment.

5.2.4 FORMULATION, HANDLING AND DISPOSAL

This type of information allows the risk assessor to determine possible pathways of release of the pesticide to the environment through activities such as mixing, loading spray equipment or discarding used containers.

These areas are most usually dealt with through standardised label statements. For example, in the case of empty containers or unused chemical:

> "Rinse containers before disposal. Add rinsings to spray tank. DO NOT dispose of undiluted chemicals on site. If not recycling, break, crush or puncture and deliver empty packaging for appropriate disposal to an approved waste management facility. DO NOT burn empty containers or product".

Within the assessment report, a statement as to the likely extent of environmental exposure due to formulation, handling and disposal should be provided. For example:

> "Formulation will not expose the Australian environment to <chemical> as it will be formulated overseas and imported as the end product. The end use product will be marketed in 1 and 5 L containers. The company does not intend to commercially repack the product in Australia.
>
> In the event of accidental spillage, it is necessary to take all possible precautions to avoid the product entering drains and waterways. Statements to this effect appear on the material safety data sheet (MSDS)."

The information provided in this section allows the assessor to determine the quantities and ways in which the product may be released to the environment.

Some use patterns will require much more assessment in the area of disposal, for example, spent dip solutions or downstream disposal of treated products such as treated timber. DEWHA has developed standard criteria for the consideration of disposal of spent dipping solutions to land, as provided in the MORAG and repeated here.

The following agreed set of criteria have been derived on the basis on analysis of the environmental fate and toxicity of 10 priority chemicals used in dips for the disposal of spent dipping solutions to land.

1. Half-life in soil <10 days at the likely concentrations following dip disposal;

   AND/OR

2. The active(s) must be able to be denatured safely, quickly and completely (>98% in 2 hours) prior to disposal.

   AND

3. If repeat applications are to be made to the same site and denaturing is not possible, these should not occur until 4 half-lives have passed.
4. The spent dip should be evenly spread over flat land at a rate not exceeding 100,000 L/ha for spent sheep dips and 20,000 L/ha for spent fruit dips.
5. The disposal site must be dedicated and adequately bunded (soil at least 15 cm high).

While DEWHA is currently undertaking an examination of the data holdings and label statements of all current active constituents and their associated products used in dipping, any application for new active constituents or extension of existing actives and associated products to be used in dips will need to be accompanied by data in the above areas to allow assessment of whether disposal to land is feasible and/or the drafting of suitable label statements for all products containing dipping uses.

5.3 ENVIRONMENTAL CHEMISTRY AND FATE

Assessing the environmental fate of a chemical provides an indication of what happens to a pesticide once in the environment, and allows the likely exposure levels for non-target organisms to be determined. The determination of the fate of the chemical takes into account its physical and chemical properties, transformation processes, mobility, and any field studies. Assessment of the data follows a multi-media approach, which allows

the behaviour of a chemical in soil, water and air, the potential for its uptake by plants or animals, and the potential for bioaccumulation in organisms to be determined.

Laboratory studies relating to environmental chemistry and fate are designed for the purpose of characterising the likely persistence and mobility of a chemical in the environment, and to identify any significant metabolites. While not formal DEWHA policy, there is a general acceptance that where metabolites exceed 10% of the initial parent concentration, they are deemed significant and must be considered separately within the risk assessment for their environmental exposure, toxicity and risk. In addition, degradates of known ecotoxicological concern should be quantified and identified even when present at less than 10% of the applied pesticide.

As discussed in Chapter 3, the types of environmental fate studies required depend on the use of the pesticide. Certain laboratory studies (eg hydrolysis, photolysis, and soil metabolism) are considered fundamental and are routinely conducted for all outdoor use pesticides. Other studies (eg photodegradation in air, volatility, and bioaccumulation) may be triggered by use/application patterns and basic physical and chemical properties of the substance. The information provided by these studies is used to determine the persistence in various environmental media, mobility, and bioconcentration potential of a pesticide active constituent and its major degradates, or of non-active ingredients in a product.

The studies outlined in Chapter 3 may be grouped in the following categories:

- **Physico-chemical degradation**: This includes hydrolysis and photodegradation in water, soil, and air. Hydrolysis studies determine the potential of the parent compound to degrade abiotically in water, while photodegradation studies determine the potential of the parent pesticide to degrade in water, soil, or air when exposed to sunlight. During these studies, data are also collected concerning the identity, formation, and persistence of breakdown products.
- **Biological degradation**: These studies include aerobic and anaerobic soil metabolism, and aerobic and anaerobic aquatic metabolism. The soil metabolism studies determine the persistence of the parent pesticide when it interacts with soil micro-organisms living under aerobic and anaerobic conditions. The aquatic metabolism studies produce similar data that are generated by pesticide interaction with micro-organisms in a water/sediment system. These studies also identify breakdown products that result from biological degradation.
- **Mobility:** These studies include leaching and adsorption/desorption, laboratory volatility, and field volatility. The leaching study assesses the mobility of the parent pesticide and degradates through columns packed with various soils. The adsorption/desorption study determines the potential of the parent pesticide and degradates to bind to soils of different types. The potential mobility of the parent pesticide and each breakdown product is determined by examining the data from both of these studies and may range from immobile to highly mobile.
- **Bioconcentration:** These studies use aquatic organisms to estimate the potential of a pesticide, under controlled laboratory conditions, to partition to the organisms from water or air through respiratory and dermal pathways. These studies also provide information on the degree to which bioconcentration of a pesticide or degradates can be reversed (depuration) should levels in the surrounding aquatic environment be reduced.

  Australia does not have formal criteria for requiring bioconcentration studies; rather, the need is determined by expert judgement. In the EU, a log Kow >3 is used as a general trigger to require a fish bioconcentration study. However, where it can be justified that exposure leading to bioconcentration is not likely to occur, a study is not necessary. Where bioconcentration is not expected because a substance is not stable in water, the study should not be required. Consequently, where the DT90 in the whole system from a water/sediment study is <10 days, a fish bioconcentration study should not be necessary, unless the proposed use of the active substance includes multiple applications at intervals short enough to result in significant long-term exposure.

  Bioconcentration is one aspect of bioaccumulation. For some chemicals, uptake through food is also important. Direct studies of biomagnification through the food chain are not usually required and are difficult to undertake. Information on modeling this pathway is provided in section 6.4.
- **Field dissipation**: Depending on the use of the pesticide, field studies may include terrestrial field dissipation, aquatic dissipation, forestry dissipation, combination products and tank mix use dissipation, and long-term field dissipation. These studies characterise the relative importance of each route of dissipation of the pesticide and its major degradates. Data generated from field studies can provide more realistic estimates of the persistence and transport of a pesticide and its breakdown products when the parent pesticide is applied under actual use conditions at representative field sites. To this end, Australia accepts international

field dissipation studies in the first instance. These are assessed for their relevance to Australian conditions and use patterns. Often, these studies may provide a more conservative result for use in Australian assessments as they are undertaken in colder climates hence resulting in longer half-lives than may be found under Australian growing conditions. Where use patterns and/or formulation types used in international studies are not reflective of those in Australia, Australian field studies may be requested.

- **Spray drift and vapour drift**: The issue of spray or vapour drift is not specific to active constituents and Australia tends not to request specific studies in this area. When considering spray drift, DEWHA assesses this in accordance with the APVMA's operating principles in relation to spray drift risk (APVMA, 2008), and this document should be consulted for a more detailed discussion on this issue. The model of choice for predicting spray drift through aerial spraying is the AGDISP model, and the APVMA sets the standard input parameters for use in the spray drift assessment. For ground application, the ground model from the AgDRIFT model is preferred by the APVMA, using standard settings. For example, high boom (unless likely use only of a low boom is clear), and the 90th percentile data for a fine to medium spray quality, and 50th percentile for medium to coarse sprays.

  Vapour drift is generally only considered in cases where the formulation is volatile, such as with some esters.

5.4 EXPOSURE CALCULATIONS

Once a risk assessor has estimated the release of a chemical, its fate in the environment and its behaviour in different environment media or compartments, then the last stage in the exposure assessment is to determine how much of it will be in each compartment, and how much organisms will be exposed to. This involves deriving predicted environmental concentrations (PECs). Organisms are predominantly exposed through direct contact, the food they eat and the water they drink. Exposure to organisms may be related directly to the application rate and corresponding drift and run-off estimates. For example, exposure to bees and non-target arthropods can be through direct overspray and therefore based on the maximum application rate, while that to non-target terrestrial plants is based on drift and run-off estimates using the maximum application rate. This is illustrated further in Section 5.4.2.

The PECs calculated in this chapter are later compared to the effects (determined in Chapter 6) in the risk characterisation stage of the risk assessment (Chapter 8).

DEWHA employs an iterative approach for risk assessment whereby a worst-case (Tier I) assessment is initially performed for each compartment. The worst-case assumptions are predominantly applied to the exposure estimation. During the risk characterisation stage (Chapter 7), if the PEC calculated using the worst-case scenario is shown to result in an acceptable risk, then no further assessment is undertaken. If the risk is unacceptable, then the exposure estimates may be refined. Such refinement options post-Tier I are discussed in this chapter at the end of each sub-section.

Several PECs require development through the exposure assessment including:

PEC_{food}	Used for assessing exposure to birds and mammals.
PEC_{soil}	Used for assessing exposure to soil organisms.
PEC_{water}	Used for assessing exposure to aquatic organisms. May also be used in the prediction of the $PEC_{sediment}$ value (see section 5.4.4.1) and for assessment of secondary exposure (see section 6.4)
$PEC_{sediment}$	Used for assessing exposure to benthic organisms.

The first two PECs predominantly relate to exposure of terrestrial organisms including plants, birds, mammals, insects, etc (Sections 5.4.1 and 5.4.2). The second two PECs are applied to the aquatic compartment and may include exposure to organisms such as fish, insects and plants (Section 5.4.4). Special- case organisms such as frogs (which may be partially aquatic and partially terrestrial) and reptiles are discussed in Section 5.4.5.

The following sub-sections address the exposure predictions for the various plant and animal communities considered during the risk assessment.

5.4.1 TERRESTRIAL ORGANISMS

5.4.1.1 Exposure estimates for birds and mammals

Dietary exposure – screening calculations

Birds and mammals are predominantly exposed to a pesticide through ingestion, either of contaminated food or water or soil, or of the pesticide itself (eg granules). Experience shows that consideration of other routes such as

inhalation and dermal contact are rarely needed. In any event, dermal and inhalation toxicity tests are not required for birds and while inhalation data are usually available for mammals, the dermal data would be of limited value as the test substance is applied to the shaved skin. Consequently, a quantitative risk assessment through the dermal and inhalation routes is not possible. The EU identifies this as a serious gap, stating efforts should be made to develop models in the near future (EC, 2002 (a)) and US EPA scientists are currently exploring methods for estimating exposure through these routes <http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_exp.htm>.

Birds and mammals may be exposed orally through food and water ingestion as well as non-food sources such as granules, baits and soil ingestion. At the Tier I level it is assumed that animals obtain all their diet from the treated area and the contaminated food is not avoided. Considerations differ depending on the use pattern, eg, sprayed formulations as opposed to granules, baits or treated seed.

Spray formulations: For spray applications, DEWHA estimates pesticide concentrations in animal food items with the focus on quantifying possible dietary ingestion of residues on vegetative matter and insects. Residue estimates are based on the updated Kenaga nomogram (Pfleeger *et al*, 1996) that relates food item residues to pesticide application rate. The Kenaga nomogram was developed by the US EPA to estimate wildlife exposure to pesticide residues on consumed plant material. The nomogram is a graphical representation of the modeled relationship between pesticide residues on the plants and the application rates and physical factors like the surface area and mass of foliage and how much of the spray is intercepted by the foliage. Residues are then compared directly with dietary toxicity data or are converted to an oral dose for use in the risk characterisation.

As an example, the following figure shows a screenshot of the calculations for residues in vegetation and other food sources based on a pesticide applied at 150 g/ha.

For a given application rate, calculate concentration in vegetation and other food sources

ENTER application rate >>>> 150 g ai/ha

Environmental Compartment	READ concentration fresh weight (mg ai/kg)	Wet/dry weight ratio (see note 2)	READ concentration dry weight (mg ai/kg)
	fresh food		dry food
short grass	32.1	3.3	106
leaves and leafy c	18.1	19	344
forage crops	18.1	5.4	98
small insects	18.1	3.8	69
grain/long gras	14.7	4.4	65
pods with seed	2.01	3.9	8
large insects	2.01	3.8	8
fruit	2.01	7.6	15

Dose to organism from residues on plants is then calculated based on the following dietary requirements:

Rats = 100% grain
Quail = 30% small insects, 70% grain (same as long grass)
Mallard = 30% grain, 70% large insects.

Therefore, residues in a quail diet for example (PECfood), are calculated as:

PEC_{food} = 0.3 ×(small insects concentration = 18.1) + 0.7× (grain/long grass concentration = 14.7) = 15.7 mg/kg fresh weight.

Mammals are generally considered on a case-by-case basis in Australian assessments. This is partially due to the lack of data on Australian species. For example, the US EPA converts the residue concentrations to a daily oral dose based on the fractions of body weight consumed daily as estimated through mammalian allometric[4] relationships. There are no consolidated statistics on Australian mammals available to DEWHA for such an exercise. The US EPA's *Wildlife Exposure Factors Handbook* may provide surrogate information for this exercise

[4] Allometry is the study of the relationships between the growth and size of one body part to the growth and size of the whole organism. Allometric relationships also exist between body size and other biological parameters (eg metabolic rate).

(US EPA, 1993). Additionally, Appendix I of EC, 2002 (a) provides daily food intake rates of some wild birds and mammals based on an extensive 2002 review. In this reference, the estimates of food intake are based on means of daily energy expenditure for free-ranging animals, energy content, moisture content and assimilation efficiencies. Given the lack of Australian data in this area, surrogate values from international sources should be used with caution.

Granular, bait and seed treatment: For granular, bait, and treated seed applications, the amount of pesticide per unit area is estimated. In this instance, direct consumption of granules, baits or treated seed by birds or mammals forms the basis of the PECfood. The label rate of application for the active constituent (ac) is the basis for the exposure calculation.

Exposed chemical per square metre can be calculated in several ways depending on whether the material is applied in rows or broadcast over the entire application site.

- **Broadcast:** This is a direct conversion. That is, 1 kg ac/ha = 100 mg ac/m2.
- **Row/band/in-furrow:** Information on the actual area within a hectare should be included. For example, if this type of application only results in half the hectare being treated, the application rate of 1 kg ac/ha results in an average 50 mg/m2. This further assumes there is no incorporation within the rows. Australia tends to assume no incorporation into soil unless the method of application suggests otherwise. In contrast, the US EPA makes the following assumptions:
  - in-furrow applications assume 1% of granules, bait, or seed are unincorporated
  - banded treatments assume 15% of granules, bait, seeds are unincorporated
  - broadcast treatment without incorporation assumes 100% of granules, bait, seeds are unincorporated.

Chapter 8 explains further how the resultant PEC values obtained through this exercise are used in the risk characterisation.

Dietary exposure – refinements when required

The above approach considers acute exposure and is reflective of the screening (Level I) methodology applied in Australia, the USA and Canada. It doesn't consider factors such as the impact of multiple applications, or short term and long-term exposure. Conversely, these aspects are considered in the EU approach even in their Tier I assessment. The source document (EC, 2002 (a)) should be consulted for a full description of the EU processes, but they are summarised below:

Vegetation following spray application: For tall growing crops (eg orchards and vines), it is assumed that 60% of the applied rate reaches the ground, which is their maximum percentage where application is made to crops without leaves. The level of interception increases at later stages of crop growth and so the amount reaching the soil decreases, meaning this is a worst case assessment.

Insects following spray application: Data for large insects are taken from a generic database originating from 24 field studies. Insectivorous mammals are assumed to consume large insects. While the values obtained from the Kenaga nomogram are used for small insects, the EU considers that research is highly desirable to develop more robust data for residues in insects, including how residue levels vary through time.

Standard exposure scenarios: The EU has developed standard exposure scenarios for a number of cropping situations as shown in the following table, *which includes default body weights* for generic indicator species. In each crop category, several indicators with different feeding preferences may be relevant. For the Tier-1 assessment, however, the number of scenarios has been restricted as far as possible. In the case of mammals, herbivorous species (if relevant) clearly represent the worst case, because, independent of their size, they receive higher doses than small omnivores (arthropods:seeds:vegetation = 1:1:1), and insectivores. With birds the situation is somewhat different; as the exposure of insectivorous birds is based on residues in small insects (as opposed to large insects with mammals). Consequently, the exposure is higher or close to herbivorous species and therefore two scenarios are proposed for some crops.

Crop	Crop stage	Indicator species	Example
Grassland	-	Small herbivorous mammal - 25 g	Vole
		Large herbivorous bird - 3000 g	Goose
		Insectivorous bird - 10 g	Wren, tit
Cereals	Early	Small herbivorous mammal - 25 g	Vole
		Large herbivorous bird - 3000 g	Goose
		Insectivorous bird - 10 g	Wren, tit
	Late	Insectivorous mammal - 10 g	Shrew
		Insectivorous bird - 10 g	Wren, tit
Leafy crops	Early / late	Medium herbivorous mammal - 3000 g	Hare
		Medium herbivorous bird - 300 g	Partridge, pigeon
		Insectivorous bird - 10 g	Wren, tit
Orchard / vine / hops	Early / late	Small herbivorous mammal - 25 g	Vole
		Insectivorous bird - 10 g	Wren, tit
Seed treatment	-	Granivorous mammal - 25 g	Wood mouse
		Granivorous bird - 15 g	Linnet

Residues in vegetation may be negligible in the case of herbicides applied to bare soil. In this case the use of herbivores as indicators may not be relevant. However, if the active substance is systemic then the risk to herbivores should be assessed. It also should be assessed as to whether there is a risk from other routes of exposure (eg soil invertebrates and earthworms), especially if the pesticide is found to be persistent.

Food intake rates for the indicator species are then calculated based on means of daily energy expenditure for free-ranging animals, energy content, moisture content and assimilation efficiencies.

Acute exposure: Consideration is given to the possible accumulation of residues on foliage due to multiple applications. A simple model based on first-order decline is used to calculate multiple application factors (MAF) that give the ratio of the initial concentration after the last of *n* applications compared to the initial concentration after the first application. MAF is a function of the number of applications, interval, and DT50. In the first tier a default value of 10 days for DT50 on vegetation is used. Special MAF factors have been calculated in order to predict the true 90th percentile of the peak after *n* applications based on the log distribution of the residue data. In the case of insects little is known on the time-course of contamination. However, it is expected that no MAF is applied for residues in insects. This is because it is anticipated that repeated applications do not cause appreciable accumulation of residues (at least in foliage dwellers) as individuals are at least partially replaced through migration and reproduction.

Short-term exposure: This assessment is only conducted for birds and considers a time frame of a few days. Consequently, initial residues are more appropriate than time-weighted averages. As usual in the first tier, animals are assumed to feed on the treated field only, however in the course of some days they will gather food in an area that is large compared to the spatial scale of residue variation. Consequently averaging of residues is expected to occur and therefore arithmetic means are taken for residues in vegetation and insects (small insects: "typical limit"). Multiple applications are again considered. However, standard MAF values are applied because residue estimates are based on arithmetic means.

Long-term exposure: The exposure estimate is very similar to the short-term assessment. Again residue estimates are based on arithmetic means, and for vegetation the same multiple application factors are employed.

In contrast to the short-term assessment where initial residues are considered, the time-weighted average (twa) residues are used here as these better reflect long-term exposure. It is obvious that a constant exposure level (if above the response threshold) will have more serious long-term effects than an exposure pattern that starts with the same level and then rapidly declines. Such long-term effects may be driven by accumulation of the substance (increase of body burden) or accumulation of effects. This has to be considered when relating toxicity (constant exposure level) to field exposure. An appropriate means to

reduce this type of bias is to average the exposure over a certain time interval. While there is no sound scientific basis and no generally accepted rule on how long this interval should be, a period of 3 weeks is proposed as a convention, unless there are good reasons to take shorter or longer times. With regard to residues on vegetation a simple twa factor is used in the first tier which is based on the following default values:

- time window (averaging time) = 3 weeks
- DT_{50}=10 days.

Using these assumptions, f_{twa} is 0.53 (ie the fraction remaining using a time weighted average approach) for residues on plants. This means that over a period of three weeks the average concentration is about half the initial concentration. (Note: In case of repeated applications the maximum twa may be underestimated when the interval between sprays is shorter than the time window; using the default values listed above the inaccuracy is small and the factor of 0.53 can be used uncorrected).

In the case of insects no default twa-factor is employed in the first tier. It is assumed that 100% of the initial residue level remains, as the variation in residue levels in these organisms through time is unknown.

Many birds are extremely mobile and therefore may potentially experience concurrent and repeated exposure in adjacent fields. This is particularly relevant for long-term assessments. In the standard procedure, the risk from multi-field scenarios is addressed by the conservative assumption that one bird obtains all of its food all of the time from the treated area. However, care has to be taken if the feeding area is refined.

Note on seed treatments: Due to the fact that there may be a long-term effect from short-term exposure, there is the need to assess the long-term risk from compounds of this type. However, this assessment is difficult because reproductive effects are only tested in studies with long exposure periods (six weeks to one year). Considerations for refinement are: availability of seeds, palatability, and degradation from seed surface. At least for some kinds of seed there may be information available on the proportion in the diet of mammals and birds. If the compound is systemic and exposure is considered likely via the consumption of treated vegetation then this should be assessed appropriately.

Exposure via drinking water – screening calculations

This is an area previously only considered for certain cropping situations in Australia (such as rice) where birds may obtain all their drinking water directly from treated bays. However, there are other avenues for birds and mammals to be exposed through water and the EU provides guidance for this route of exposure (EC, 2002 (a)).

Species that frequent open water bodies (eg water birds or migratory species) are liable to ingest residues of active substances that reach water, such as by spray drift from treated fields. The exposure concentration in this case is equal to $PEC_{surface\ water}$.

In some situations, some species may obtain all their daily water demand directly from puddles of spray liquid or reservoirs held in the axils of leaves. This situation may be relevant for certain crops (eg vegetables) or growth stages and certain seasons (summer). This route can be considered less relevant for substances that are volatile and rapidly photolysed in sunlight. The exposure concentration can be calculated from the dilution used to prepare the product for spraying. Analysis has shown that initial concentrations in such sources are in the range 5-20% of the sprayed concentration, therefore, a dilution factor of 5 is applied (EPPO 1994).

The daily water intake is calculated allometrically as follows:

- birds: total water ingestion rate (L/day) = 0.059W0.67
- mammals: total water ingestion rate (L/day) = 0.099W0.90

where W is the body weight in kg.

Thus, the daily dose of active substance is calculated as ($PEC_{drinking\ water}$ × total water ingestion rate)/W.

Refinement options

If these Tier I calculations for exposure via diet and drinking water indicate that risk to birds and mammals may be unacceptable, then several options exist to refine the exposure assessment, as detailed below.

Measured residues: Residues determined by the Kenaga nomogram in the first tier are based on generic data. Refinement may be possible by making use of available residue data for the substance and conditions to be assessed.

If residue trials involve repeated applications of the product and sampling starts at the last application, then the total residues from all applications are already included. This negates the need to consider a multiple application factor, as used in the EU approach.

Residue decline in plants: Previous assessment experience indicates that the disappearance of residues from plant material is fairly rapid even in the case where the chemical is persistent in other environmental media. There are different routes of disappearance of a substance from vegetation including volatilisation, wash off, degradation and metabolism.

In addition there is a decline of residues due to dilution by growth. The integrated result of these processes is usually expressed as an initial rapid decline in surface residues followed by a slower phase, so the assumption of first-order kinetics may be inappropriate when long time frames are considered. Useful information may be derived either from a general database or from the substance under assessment (EC, 2002 (a)).

Avoidance: Avoidance may be a significant factor in reducing exposure. While palatability tests may not often be available, some indication of avoidance may be obtained from dietary studies.

De-husking: De-husking may reduce exposure of seed eating birds and mammals. Regardless, whether seed treatment is the intended use of the product or weed seeds are contaminated during spraying, the substance will be mainly on the husk and therefore dehusking can remove the majority of the residue. This reduction can be as high as 85%. Small birds are more likely to de-husk seeds than large birds, but in any event, it depends on the kind of seed, and even when de-husking occurs, as only a proportion of seeds are de-husked (EC, 2002 (a)).

Refinement to diet characteristics: The Tier I assessment assumes all diet is taken from the treated area (probably leading to an overestimate of exposure). Additionally, the Kenaga nomogram makes basic assumptions about the composition of diet that may be refined. However, in order to refine either of these assumptions, data should be available to support the refinements, such as observations to demonstrate the time that birds and mammals actually spend in treated areas.

5.4.1.2 Exposure estimates for bees and other terrestrial arthropods

Bees and many other non-target terrestrial arthropods are regarded as beneficial insects, as they are crucial for pollination, predation on other crop pests and the like.

Bees

Screening calculation

Exposure to bees is determined for spray applications based on the maximum application rate. This rate is converted to a rate of chemical (ac) per square centimetre on the assumption that a honeybee is approximately 1 cm2 in surface area (Davis and Williams 1990).

For example, a chemical with an application rate of 150 g ac/ha results in a rate of 1.5 μg/cm2 (150 g = 1.5 ×108 μg, 1 ha = 1 × 108 cm2). This results in an assumed exposure of 1.5 μg/bee.

EC, 2002 (b) states that for systemic plant protection products, exposure considerations and calculations should also include the active constituent (or metabolite) present in the respective plant parts (eg nectar and pollen) to which honeybees could be exposed. However, it should be noted that estimates of these concentrations are rarely available.

Refinement options

If predicted effects to honeybees are considered as not acceptable, a refinement of exposure may be considered through various aspects of the use pattern. For example:

- varying the application rate
- timing of application and appropriate label statements, for example:
  - DO NOT apply before 7 pm and/or during honeybee flight
  - DO NOT apply when crops are flowering.

Other arthropods

As with bees, the exposure to other arthropods (eg predators, parasites and ground dwelling organisms) is determined for spray applications based on the maximum application rate in kg/ha. Exposure is considered for both in-field exposure (direct application), and off-field exposure (spray drift). Refinements may be made to the exposure based on crop interception or taking into consideration the use pattern. For example, where application only occurs to part of the field (spot treatments or row/furrow treatments), significant areas of untreated land remain as refuges for certain non-target arthropods.

This approach cannot be compared to the USA where, currently, the US EPA does not characterise residue exposure for honey bees and other beneficial insects <http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_exp.htm#HDPDITE>. Similarly, Canada appears to make a statement as to the hazard of chemicals to terrestrial arthropods but does not appear to characterise exposure.

The EU does provide guidance for estimated exposure to terrestrial arthropods, and two source documents are available (EC, 2002 (b) and Candolfi *et al*, 2001). Within this scheme, two scenarios are considered, namely, in-field and off-field.

If exposure to other arthropods is relevant in assessing a particular pesticide, the EU guidance is appropriate and is summarised below.

Screening calculation (also known as Tier I assessment)

For the standard assessment the following scenarios are used to describe the exposure in-field and off-field. For both, the key input is the nominal field application rate (in g/ha or mL/ha) supplemented by various factors:

- in-field exposure = Application rate × MAF
- off-field exposure = Application rate × MAF * (drift factor/vegetation distribution factor)

  where MAF = multiple application factor

For calculation of MAF values, definitions and further details see Candolfi *et al*, 2001. With regard to the vegetation distribution factor, Candolfi *et al*, 2001 gives a default value of 10. However, this figure is considered unreliable and therefore when more appropriate data is available it should be brought into use in Australian assessments. The tables published by Rautmann *et al*, 2001 are used for the drift factor; the standard assessment should be conducted for 1 m distance (arable crops) or 3 m (orchards and vineyards); drift factor = % drift/100.

Refinement options

In this case refined assessments can only be based on the outcome of higher-tier studies. In such studies relevant exposure issues are considered in the study design when establishing the dosing regime (be it dose/response design or single-dose design). That makes a separate exposure assessment unnecessary. It must, of course, be ensured that the study covers the use scenario under assessment (EC, 2002 (b)).

5.4.1.3 Exposure estimates for soil organisms

Initial calculations

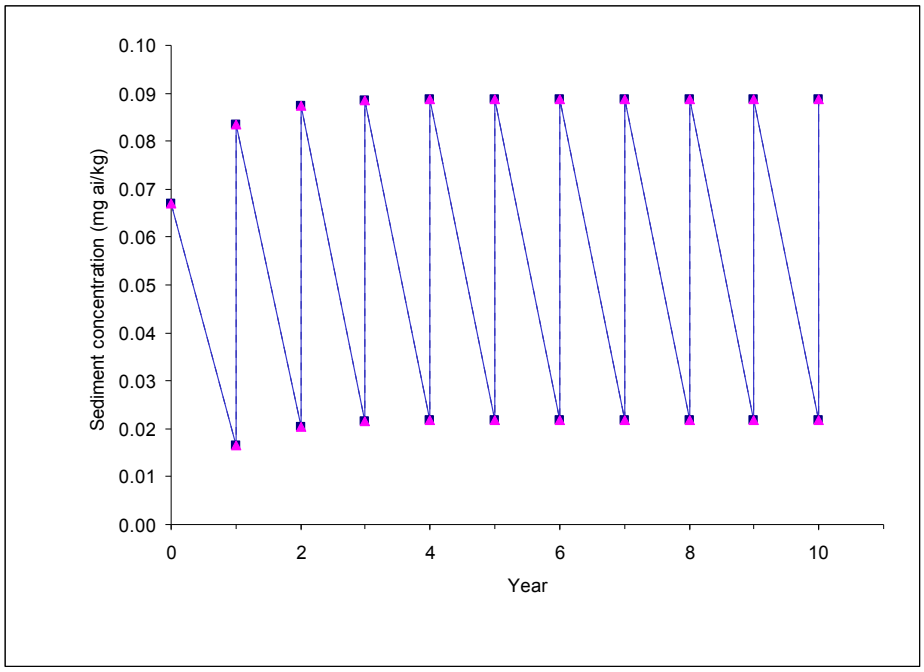
The exposure calculations for soil organisms such as earthworms and soil micro-organisms are based on the application rate of the chemical. Three situations are considered:

1. Where a single application is made, the PECsoil is based on the application rate
2. Where multiple application rates are made, the PECsoil is based on the combined rates assuming no degradation or dissipation
3. Where the chemical is considered persistent to the extent that carryover may occur between treatments or years, accumulation should be considered and the plateau concentration used for determining the PECsoil.

With regard to situations in this third category, the concentration in soil should be predicted based on uniform mixing within the top 15 cm using a soil density of 1,500 kg/m3 following the process discussed below. This is the default soil density value in the Mackay Level III fugacity model and is currently used in industrial chemicals assessments in Australia. For further rationale on this point, refer to the risk assessment manual for industrial chemicals (Lee-Steere, 2005).

To determine accumulation of a chemical in soil, DEWHA uses a process based on Smith, 1982. However, a soil half-life will be required in order to use this guidance. Assessors should determine an appropriate half-life based on available laboratory and field data, or modelling as appropriate. The concentration of the chemical in soil will not be constant in time. For example, Figure 5.1 below shows accumulation for a chemical applied to soil at a rate of 150 g/ha with a half-life of 180 days in soil:

Figure 5-1: Soil accumulation in top 15 cm following 10 years of pesticide application at 150 g ac/ha.



Annual carry over of the chemical is determined by the slope of the regression equation, or the removal rate constant, (Slope = ln(2)/half-life) and is calculated as follows:

$$\text{Annual Carry over (\%)} = e^{\text{slope}.365} \times 100$$

From the annual carry over, the concentration immediately following application in future years can be calculated. The accumulation value immediately following application should be used for the PECsoil in the first instance for situations as described in dot point 3. Where secondary exposure is being considered in such situations (see Section 6.4), it is more appropriate to use the concentration in soil half a year following application, as this will approximate an annual average.

Refinement of PEC_{soil}

Degradation: Where multiple applications occur and reliable soil degradation data exist, this information may be used to refine the PECsoil. For example, if the application rate entails four consecutive sprays 14 days apart, a representative half-life may be used to predict soil concentrations immediately following each application. Where chemicals are persistent, this may not prove a good refinement technique and assessors should use the process described above. However, where chemicals do not persist, the PECsoil may be significantly reduced.

Crop interception: A further way of refining the PECsoil is to consider removal during application as a result of crop interception. The EU considers this in a rudimentary fashion within the screening level PEC calculation. In their process, unless better information is available the fraction intercepted is assumed to be 0 for applications to bare soil, or up to 0.5 for applications when a crop is present (Boesten *et al*, 1997). This issue is noted as one requiring much more consideration. Australia does not currently have a standardized approach for considering crop interception but it is included where required on a case by case basis. Foliar interception rates for select crops are published in the literature, and these may be used as guidance (Linders *et al*, 2000).

Field studies: Data from relevant field studies submitted by the registrant can be used where available. Boesten *et al*, 1997 provides useful guidance on the use of field dissipation data in refining the PECsoil:

> For the first stage assessment of PEC values in soil it is appropriate to use laboratory transformation rates. However, field dissipation studies do have an important role to play, since they are a direct measure of concentrations in soil under field conditions. Field dissipation studies should therefore be considered more definitive than modeling predictions if conducted under relevant conditions. However, the use of field data needs care to ensure that the results have predictive value. For example, if leaching of a highly mobile chemical significantly speeds dissipation on a very sandy soil, the resulting dissipation rate has limited value as an indicator of what will happen on fine-textured soils. Another restriction with respect to the use of field data is the soil temperature.
>
> If field dissipation studies have been carried out at temperatures that are, for example, 10°C warmer than relevant, this may lead to interpretation problems. If field dissipation is significantly faster than aerobic soil

transformation rates in the lab, this may indicate that additional processes may be important in the field, eg photolysis, volatilisation. If this conclusion is supported by data specific to these processes (eg lab photolysis rates, vapour pressure) then it may be appropriate to use the field dissipation rates to refine the estimation of PEC values in soil. It is sometimes possible to support the use of the field data by eliminating the possibility that other dissipation routes were important in the field studies (eg eliminating run-off if the site was flat). Ideally, it may be possible to use detailed simulation models to show that the field results are consistent with lab-based estimates of rates of transformation, photolysis etc. However, this is often not possible, either because of incomplete characterisation of the field studies, or because of the inadequacy of some existing lab methods (some lab study designs are designed to demonstrate the existence of particular dissipation routes, not to produce numbers which can be used to predict field rates of the same process).

In any case, if field dissipation data exist, then they should be considered in the assessment. Field dissipation rates often differ from laboratory transformation rates, and the reasons are not always clear. The possibility that some laboratory studies are not predictive should also be considered, because of the difficulties of maintaining active micro-organisms in the lab over long periods of time, and inadequacy of protocols followed, often decades ago. It could be argued that there is less to go wrong with field studies provided that an adequate soil sampling strategy is applied. Field dissipation studies also have particular value when the rate of transformation in the laboratory is too slow to determine with certainty.

The EU uses more complex modeling to refine soil PEC values, and a full discussion on the model types may be found in the source document on soil persistence models and EU registration, Boesten *et al*, 1997. Prior to using models referred to in this document for refining soil exposure concentration estimations in Australian assessments, the models should be assessed for their applicability for use in Australia, including country/region specific input requirements.

5.4.2 EXPOSURE ESTIMATES FOR NON-TARGET TERRESTRIAL PLANTS

Currently, different countries differ in their approach to assessing risks to non-target plants. Consequently, it is not possible to define an international best practice methodology.

To this end, an international workshop was held 15-17 January, 2002 in Arlington, Virginia, to explore various approaches to assessing the ecological risk of chemicals to non-target plants. The workshop was convened by the Office of Pesticide Programs (OPP) and the Office of Pollution Prevention and Toxics (OPPT) in the US Environmental Protection Agency (EPA).

The purpose and goals of the workshop were as follows:

- improve communication and collaboration among the international community regarding regulatory approaches for conducting risk assessments for non-target plants
- discuss the elements of a risk assessment framework and the issues associated with exposure and ecological assessments for non-target plant risk assessments
- identify and prioritise research topics which will address the uncertainties in non-target plant risk assessments
- develop a set of recommendations for advancing non-target plant risk assessments.

The key findings of this workshop with regard to exposure assessment were as follows:

Different countries currently include different exposure pathways in exposure assessment. Aerial applications are not common in some parts of Europe, while tile drains are not found extensively in North America or Australia. Most countries do not evaluate the atmospheric pathway (volatilisation and transport of soil/dust). An effort is underway in Europe to develop ten or more defined scenarios for drift, run-off and drainage based upon GIS data for topography and climate.

Information used in exposure assessment includes the physical and chemical properties of the pesticide, information about persistence and mobility, and environmental data. For multiple applications, information on degradation, the number of applications, and interval between applications is needed. The method of application is also important. Exposure assessments typically use maximum exposure scenarios and generic information at lower tiers of the risk assessment and are more realistic (more or better-defined scenarios) at higher tiers. Uncertainty is expressed by modelling a range of scenarios using a range of input parameters.

As well, the form of the chemical that should be tested (technical active constituent or formulated product) may be important, since terrestrial plants are typically exposed to the formulated product. It was noted that exposure information is often generated based upon the technical active constituent rather than

the formulated product. It may be appropriate to develop data for the active constituent and then develop bridging data for formulation(s).

In general, a better understanding of the movement of pesticides in the environment is needed. The most important research needs for improvement of exposure assessment include: aquatic dissipation studies under varying scenarios; investigations on multiple exposures, chronic exposures, long-distance transport; applications of existing models and/or development of new models; research on plant transpiration; and improved detection methods for low-dose products.

Australia assesses exposure to non-target plants based on spray drift and run-off although a standardised approach has not been finalised. The following process describes a method for assessing exposure to non-target plants and is based on US EPA methodology:
<http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_exp.htm>

For terrestrial and semi-aquatic plants, run-off and spray drift scenarios that are based on a pesticide's water solubility and the amount of pesticide on the soil surface and its top 2.5 cm layer are used.

- Runoff: It is assumed that there is sheet run-off from one-treated hectare to an adjacent hectare for dry areas, while channelised run-off is from 10-treated hectares to a distant low-lying hectare for semi-aquatic areas.
- Spray drift: Exposure to plants from ground application is assumed to be 1% of the application rate and 5% for aerial, air blast, forced-air, and chemigation application.

The formulae for calculating PECs for unincorporated ground applications, incorporated ground application, and aerial, air blast, forced-air, and chemigation applications are listed below:

Terrestrial plants inhabiting dry areas adjacent to treatment sites

Unincorporated ground application

Run-off = maximum application rate (kg ac/ha) × run-off value (see note below on how this is calculated)

Drift = maximum application rate × 0.01 (ie 1% drift)

Total loading (or exposure rate) = run-off (kg ac/ha) + drift (kg ac/ha)

Incorporated ground application

Run-off = [maximum application rate (kg ac/ha) ÷ minimum incorporation depth (cm.)] × run-off value (see note)

Drift = drift is not calculated if the product is incorporated at the time of application.

Total loading = run-off (kg ac/ha) + drift (kg ac/ha)

Aerial, air blast, forced-air, and chemigation application

Run-off = maximum application rate (kg ac/ha) × 0.6

(60% application efficiency assumed) × run-off value (see note)

Drift = maximum application rate (kg ac/ha) × 0.05 (ie 5% drift)

Total loading = run-off (kg ac/ha) + drift (kg ac/ha)

Terrestrial plants inhabiting semi-aquatic low-lying areas

Unincorporated ground application

Run-off = maximum application rate (kg ac/ha) × run-off value (see note) x 10 ha

Drift = maximum application rate × 0.01 (ie 1% drift)

Total loading = run-off (kg ac/ha) + drift (kg ac/ha)

Incorporated ground application

Run-off = [maximum application rate (kg ac/ha)/minimum incorporation depth (cm)] × run-off value (see note) × 10 ha

Drift = drift is not calculated if the product is incorporated at the time of application.

Total loading = run-off (kg ac/ha) + drift (kg ac/ha)

Aerial, air blast, and forced-air applications

Run-off = maximum application rate (kg ac/ha) × 0.6 (60% application efficiency assumed) x run-off value × 10 ha

Drift = maximum application rate (kg ac/ha) × 0.05 (ie 5% drift)

Total loading = run-off (kg ac/ha) + drift (kg ac/ha)

The above calculations are undertaken within the US EPA GENEEC model (see Appendix II).

Note: Run-off values appear to be calculated within this model. Therefore, prior to using this approach in Australia, this model would either need to be evaluated for applicability to Australian conditions, or work undertaken on developing appropriate run-off values. For information on run-off values in the above context, consult <http://www.epa.gov/oppefed1/ecorisk/rra_chap_four.htm>. It is expected further work on this issue could be considered within the context of NChEM.

5.4.3 GROUNDWATER

Groundwater concentration is an area considered on a case-by-case basis in Australian assessments. However, it may be calculated for determining indirect exposure of humans through drinking water, and can be predicted using the approach described in the TGD. This methodology uses the concentration in porewater of agricultural soil as an indication for potential groundwater levels. It is a worst-case assumption that neglects transformation and dilution in deeper soil layers.

The formula used to calculate the groundwater concentration is essentially:

$PECgw = PECsoil\ porew = (PECsoil \times BDsoil)/(Kpsoil\text{-}water \times 1000)$

Symbols			
Parameter	**Symbol**	**Unit**	**Value**
Local PEC in groundwater	PEC_{gw}	mg/L	
Local PEC in soil porewater	$PEC_{soil\ porew}$	mg/L	
Local PEC in soil	PEC_{soil}	mg/kg	
Soil-water partition coefficient	K_p	m^3/m^3	
Bulk density of soil	BD_{soil}	kg/m^3	1500 kg/m^3 See below

The bulk density of soil is determined by considering the fraction of solids, water and air all multiplied by their respective densities. There are currently no Australian specific default values for these respective parameters. The Level III fugacity model has default values of 20, 30 and 50% air, water and solids respectively for the soil compartment, with respective densities of 1.21 kg/m3, 1000 kg/m3 and 2,400 kg/m3 giving a default soil bulk density of 1500 kg/m3. Until Australian specific values are obtained, it is recommended the value from the Level III model be used to maintain consistency within the exposure assessment. It is expected further work on this issue could be considered within the context of NChEM.

5.4.4 EXPOSURE ESTIMATES FOR AQUATIC ORGANISMS

5.4.4.1 Calculations for screening level PECwater

As noted in section 5.4, when assessing risk, it is not possible to account for every scenario and case. Consequently, Australia follows an iterative process by considering:

- a 'worst case' exposure scenario, and, if needed
- a series of refinements which account for other factors and results in setting more realistic scenarios at each step.

The worst-case scenarios considered for exposure of aquatic organisms are that of a direct over spray and a 10% spray drift to a standing body of water 15 cm deep with a 1 ha surface area (ie 1.5 ML of water). The direct over spray concentration is also equivalent to a worst-case of 10% run-off from a 10 ha watershed to the same size standing body of water. The number of consecutive applications within a crop cycle should be considered on the assumption there is no degradation or dissipation from the water column. Therefore, the screening level PECwater will be defined in terms of active constituent as:

$PECwater\ (mg/L) = (Rate\ (kg) \times 106\ mg/kg \times no\ of\ applications)/1.5 \times 106\ L$

Once a chemical is present in the water column it can partition onto suspended sediment which then settles onto the floor of the standing body of water as part of the sediment compartment. How much sorbs to the suspended sediment and how much stays dissolved in the water depends on the characteristics of the chemical. Therefore, for organic chemicals the potential concentration in the sediments can be modelled from the estimated water concentration using the equilibrium partitioning method.

For a fuller explanation of this, refer to the Risk Assessment Manual for industrial chemicals.

Kp is the solid-water partition coefficient in sediment (L/kg). This value represents the concentration of the substance sorbed to solids (mg/kg) suspended in the water column divided by the concentration dissolved in water (mg/L), and is calculated according to:

$Kp = Koc \times foc$

where,

K_{oc} = organic carbon normalized distribution coefficient (L/kg)

f_{oc} = fraction of organic carbon in sediment. An appropriate default value for f_{oc} is 0.05 (5%)

For the whole sediment compartment, the Kp is converted to account for the make up of the sediment (assumed 80% water and 20% solids based on the Mackay Level III Fugacity Model). The whole sediment compartment-pore water partition coefficient is unitless and is the concentration in solids (mg/m3) divided by the concentration in water (mg/m3), calculated as follows:

$Ksediment\text{-}water = 0.8 + (0.2 \times Kp/1000) \times BDsolid$

where,

BDsolid = the bulk density of the solid phase only, or 2400 kg/m3 using default values from the Level III Fugacity model

The final step is to convert the PECwater to a PECsediment based on the Ksediment-water partition coefficient, and the density of the bulk sediment compartment. Using Level III Fugacity default values for sediment, the density of sediment is calculated to be 1280 kg/m3 (80% water at 1000 kg/m3 and 20% solids at 2400 kg/m3).

The equation below determines the final concentration in the bulk sediment compartment accounting for that sorbed to the solids phase and that in pore water.

$PECsediment = Ksediment\text{-}water/BDsediment \times 1000 \times PECwater$

Parameter	Symbol	Unit	Value
Local PEC in sediment	$PEC_{sediment}$	mg/kg	
Sediment matter-water partition coefficient	$Kp_{sediment\text{-}water}$	m^3/m^3	
Bulk density of sediment	$BD_{sediment}$	kg/m^3	1280 kg/m^3
Local PEC in water	PEC_{river}	mg/L	Determine above

There are no Australian-specific default values. The Level III fugacity model uses a default value for density of sediment as 2,400 kg/m3 and a fraction of organic carbon in suspended sediment of 20%. Although this value of organic carbon is fairly high compared to most Australian soils, it is recommended the value from the Level III model be used to maintain consistency within the exposure assessment until Australian-specific values are obtained. It is expected further work on this issue could be considered within the context of NChEM.

Highly adsorptive substances may not be considered adequately with this approach, as they are often not in equilibrium distribution between water and solids due to their cohesion to the suspended matter and therefore the assumptions built into the equilibrium partitioning model do not apply. However, the chemicals may be desorbed after ingestion by benthic organisms so they may still be bioavailable. Therefore, this calculation may underestimate the sediment concentration when release to surface water occurs predominantly as particles. If this is expected, it should be considered in further evaluation, for example, when comparing the PEC with monitoring data for existing chemicals, and in the risk characterisation (EC, 2003 (a)).

5.4.4.2 Refinement options

The PECwater and hence the PECsediment may be refined through limiting exposure to surface water resulting from either spray drift or run-off.

Spray drift

Spray drift may occur from either aerial or ground application and both are considered separately. In addition, information may be gleaned from laboratory and field data, for example, degradation and movement from water to sediment observed in water/sediment tests.

Spray drift from aerial application

The APVMA has set the methodology for spray drift risk assessment in their operating principles in relation to spray drift risk (APVMA, 2008). In so doing, they have provided DEWHA with standardised input parameters for aerial applicaton using both fixed wing and rotary wing aircraft. Spray drift is modelled for both types of

aircraft at maximum wind speeds of 8, 15 and 20 kph. Spray droplet size is based on the band immediately below that being used. For example, where a coarse droplet size is advised, modelling will be undertaken using a medium-coarse spray. The model to be used for aerial drift assessment is AGDISP.

DEWHA models output on a 3 m wide water body with a depth of 15 cm. This is considered conservatively representative of Australian conditions in cotton growing areas (broad-acre crop).

PEC_{water} values are calculated by the model at various buffer distances (from the edge of the crop field). From these values, revised $PEC_{sediment}$ values at the corresponding buffer distances can be calculated as described above. Alternatively, where the eco-toxicity assessment end-point is known, this value can be used directly for the model to provide the appropriate buffer zone.

Spray drift from ground application

In their operating principles, the APVMA (2008) makes the following comments with respect to ground application:

> The modelling components of AGDISP and AgDRIFT for aerial application are very reliable and well validated, but AgDRIFT lacks a true ground modelling component, and the ground model in AGDISP has not been validated against field data and sometimes behaves erratically. Therefore, for ground application spray drift risk assessment, the APVMA relies on available field data that have been collected in well-controlled studies. The APVMA uses data sets originating in the USA, Canada and Germany. For most situations, the North American data more closely match Australian conditions. The North American studies are also more comprehensive and well validated.

Previously, the APVMA has used a data set originating in Germany (Ganzelmeier *et al*., 1995) and revised by Rautmann *et al.* (2001) to model the risk associated with ground and orchard applications. However, the field studies for the revised Ganzelmeier tables for spray drift on broad acre, orchard and vegetable crops were conducted under ideal conditions in Germany with low wind speeds, with optimally calibrated spray equipment and importantly delivering a coarser droplet spectrum than is appropriate for this assessment. The larger droplet spectrum in particular accounts for most of the much lower spray drift risk indicated by the Ganzelmeier tables for boomspray application. For most Australian assessments, DEWHA now considers that the data sets incorporated into the AgDrift ground and orchard models (AgDRIFT Spray Drift Task Force Spray Software, Version 2.0.09) are more representative of the diversity of conditions and practices undertaken in Australia. Consistent with current APVMA policy (APVMA 2008), ground application is therefore assessed using the AgDRIFT Tier I ground application model. For broadacre agriculture application, a high boom setting is used with 90th percentile data for fine to medium spray and 50th percentile data for medium to coarse spray. These may be amended for other uses. For example, turf application may use a low boom and 50th percentile data.

The corresponding refined $PEC_{sediment}$ can be calculated from the revised PEC_{water} as described above.

Observations from laboratory and field data

Where valid data exist on information such as dissipation from the water column (water/sediment studies), these data may be used to further refine the exposure. For example, where multiple applications occur, degradation/dissipation from the water between applications should be considered. Where this refinement approach is used, care should be taken in predicting the corresponding $PEC_{sediment}$ as removal from the water column may be due to partitioning to the sediment and not degradation.

Run-off

The worst-case calculations described above use a 10% run-off rate from a 10 ha field (depending on the crop) into a 1 ha standing body of water 15 cm deep. This equates to a direct over-spray situation and usually will result in an unacceptable risk.

Because actual data on run-off are rare, Australia may use modelling in an attempt to quantify run-off. However, this has not been done in a consistent manner in the past. The model described by Birkved and Hauschild, 2003 has been used in the past as it considers the Koc of the chemical and calculates the fraction of applied chemical likely to run-off with consideration of soil type (based on sand content), slope of fields and rainfall. An additional calculation is made for the dissolved fraction of pesticide in run-off waters. Another published model, Mensink *et al*, 1996 has also been used. This model accounts for factors such as the Kow of the chemical and soil organic carbon and provides a method to semi-quantitatively determine the amount of pesticide in surface water from run-off. Both models may be run in an assessment, and the strengths, weaknesses and applicability of the model results are considered on a case-by-case basis.

Currently, DEWHA are reassessing their runoff models and are examining the applicability of several models. Particular attention is being paid to the Birkved and Hauschild model and conceptually similar models.

5.4.4.3 Other international methodologies

While the broad process of aquatic exposure assessment described above is in general agreement with international approaches, there are some notable differences. The approaches used by the US EPA and the EU are provided at Appendix II.

Both the US EPA and the EU use a tiered approach in their aquatic exposure assessment. However, both also take into account certain refinements within the first tier. For example, within their GENEEC 2 model, the US EPA accounts for factors such as spray drift, run-off, degradation and chemical properties such as Koc and Kd within this screening approach. Nevertheless, this approach is considered a screening level assessment as it accounts for a generic region in terms of climate, soils, topography or crop in estimating potential pesticide exposure.

If a more refined risk assessment is needed, a higher tiered screening model (eg PRZM-EXAMS) is used to estimate pesticide concentrations that are more reflective of actual use site conditions. A detailed description of these aquatic models can be found at US EPA's website at www.epa.gov/oppefed1/models/water/index.htm.

For existing chemicals, the US EPA uses PRZM_EXAMS when reliable surface water monitoring data are available, to help characterise the levels of pesticide that are being detected in the environment.

The EU uses a four-step process (EC, 2002 (c)). In this process, the first step accounts for partitioning between water and sediment based on the chemicals Koc. If the first step results in unacceptable exposure, the second step factors in degradation data, crop interception and spray drift. This second step is still considered a generic assessment.

If the third step is required, the use of higher tier exposure models (including PRZM) is used.

5.4.5 EXPOSURE ESTIMATES FOR AMPHIBIANS AND REPTILES

It is very unusual to receive toxicity data for amphibians and reptiles, and these organisms are not routinely considered in Australian assessments. The US EPA generally uses the PEC values for fish or aquatic invertebrates when required to assess exposure of amphibians, while exposure patterns for reptiles are generally considered to be comparable to birds. However, it is noted that exceptions may occur with certain aquatic organisms that lay eggs in terrestrial areas <http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_exp.htm>. In general, the processes described above would be used to calculate the environmental concentrations from use of a pesticide in the compartments relevant to amphibians or reptiles.

Amphibians have been included in some of the proposed tiered screening tests for endocrine disruption. Consequently, data on amphibians may become more common in the future.

CHAPTER 6 – ENVIRONMENTAL EFFECTS ASSESSMENT

6.1 INTRODUCTION

The assessment of environmental effects considers ecotoxicity data to determine the hazards posed by a chemical to non-target plants and animals, both terrestrial and aquatic. It involves evaluating data on lethal and sub-lethal effects in acute and chronic toxicity laboratory tests on a selected range of standard-test species including birds, earthworms, non-target terrestrial arthropods, terrestrial plants, and three trophic levels of aquatic organisms (fish, aquatic invertebrates and algae/aquatic plants).

As noted in MORAG, there are emerging areas where environmental effects data are of interest. These include combination toxicity data for formulations containing two or more active constituents, to allow assessment of any increased toxicity from the combination product; and sediment testing, which is also a relatively new emerging and important area of toxicity in the aquatic environment and is particularly important for insoluble persistent pesticides. Effects data will then be compared to the PECs calculated in Chapter 5, during the risk characterisation stage (as discussed in Chapter 8). To a certain degree, this comparison may involve an iterative process, where effects are compared to worst case exposure estimates. Unacceptable levels of risk may then require the risk assessor to refine the assessment, such as by requesting further data. In this case, the risk assessment would begin again in the effects assessment stage.

This chapter discusses assessment of effects for several compartments, namely:

- terrestrial, including sediment (Section 6.2)
- aquatic (Section 6.3)
- secondary effects, for bioaccumulative substances (Section 6.4).

Assessment of atmospheric effects is generally limited to abiotic effects such as long-range transport potential, global warming potential and ozone depletion potential as biotic effects are seldom available where exposure is through the gas phase. This has been discussed in the corresponding environmental risk assessment guidance manual for industrial chemicals and assessors should consult this document where necessary.

Several concepts should be considered prior to the assessment of effects for the different compartments (Sections 6.2-6.4). Such comment on evaluation of data, use of assessment factors, NOECs, emerging issues related to increased understanding of specific toxicity mechanisms (eg endocrine disruption potential) and formulations are contained in the following sub-Sections 6.1.1, 6.1.2, 6.1.3, 6.1.4 and 6.1.5, respectively.

6.1.1 COMMENT ON EVALUATION OF DATA

As always, the first step of the effects evaluation process should be an evaluation of available ecotoxicity data for adequacy and completeness. Evaluation of data has been discussed in Chapter 4. However, further to this, several concepts relating to aquatic toxicity should be taken into consideration in determining their adequacy (refer Section 6.3.2). Australia accepts tests performed to accepted international standards, and on internationally recognised species. It is unusual to receive test data on Australian species. Except in certain special cases, international species are taken as being representative of Australian species, especially in the context of the conservative screening approach used for assessment.

The GHS provides a discussion on concepts relating to testing and data interpretation of difficult to test substances, a summary of which is included in Appendix III. Assessors should understand these concepts prior to undertaking any review of the adequacy of aquatic ecotoxicological data.

In addition, the following chapter considers effects data received with a data package, and how results in one area (including fate and risk characterisation outcomes) may impact on requirements for other effects data.

6.1.2 COMMENT ON USE OF ASSESSMENT FACTORS

Data considered during the effects assessment may be used to determine the predicted no effect concentration (PNEC), generally through assessing the dose/response of the chemical. The term PNEC has generally not been used in agricultural chemical assessments in Australia. Where industrial chemicals are assessed for their environmental effects, a PNEC is derived as being representative of a whole environmental compartment and is based on the lowest valid effects data for a compartment with the use of assessment factors depending on the quality and quantity of data available (see Lee-Steere, 2005).

In agricultural chemical assessments, the effects data used in the risk characterisation are taken directly from the laboratory or field toxicity data, and organisms within environmental compartments are considered separately in Australian assessments. Assessment factors are not assigned to data, rather they are implicitly built in to the risk

quotients obtained by comparing exposure (PEC) to the relevant effects data. As a result, consideration may be given to the need for mitigation of risks when quotients are above a range including 0.1 to 0.5 depending on the type of organism and type of data available. For industrial chemicals such considerations are undertaken when quotients are above 1.

This is consistent with international assessment schemes. The US EPA use risk quotients (exposure/toxicity) with level of concern values serving as triggers to identify whether further refinement of assumptions or extra data are required. The EU predominantly uses toxicity exposure ratios (TERs = toxicity/exposure). These values are the reciprocal of the risk quotient.

While the approaches are similar (and thus allow work-sharing amongst different countries), the ratios deemed as levels of concern differ amongst jurisdictions. That is to say, the in-built assessment factors are different. There is not yet any harmonisation amongst OECD member countries in regards to the in-built assessment factors. These differences are highlighted in Section 8.8.

6.1.3 COMMENT ON USE OF NOECS

The concentration at which 50% mortality occurs is defined as the median lethal concentration (LC50). The concentration at which there is no observed adverse effect is defined as the *no observed effect concentration* (NOEC). For non-aquatic scenarios, a *no observed effect level* (NOEL) is used.

The DEWHA, along with other international assessment agencies, use long-term NOECs for characterisation of environmental effects. One aim of the ecological risk assessment is to predict effects on the population level, although this is difficult or impossible to measure directly. The usual approach is based on the consideration that effects on populations will not occur if the survival rate, reproduction rate and development of individuals are not affected. Therefore, in principle, only end-points in toxicity tests that are related to these key factors of population dynamics are ecotoxicologically relevant. By definition the NOEL is based on the most sensitive end-point of the test, and that is used in Tier I. In a refined assessment it could become necessary to check the ecological relevance of the effects seen at doses above the NOEL. For example, reproduction tests with mammals and birds include parental and reproductive end-points. If the overall NOEL is based on a reproductive end-point but exposure will be transiently outside the breeding season then the NOEL for parental effects would be more relevant (EC, 2002 (a)).

It is important to note that use of NOECs is under development and revision. OECD(1998) concluded that the NOEC, as the main summary parameter of aquatic ecotoxicity tests, is inappropriate for a number of reasons, (see detailed discussion in the source document) including that the NOEC is set by the study design rather than a statistical evaluation of the results and should therefore be phased out. OECD, 1998 recommended that the OECD should move towards a regression-based estimation procedure. The time course of effects should be incorporated in the analytical procedures, and the OECD should initiate a study of the available time-dependent regression models (both mechanistic and empirical) in order to select those which best meet its needs. This study should also address the issue of appropriate values of x for ECx and the optimal experimental designs. Although there has been work done in this area, consensus amongst OECD countries has not yet been reached.

Since then, the OECD has released a draft report on the statistical analysis of ecotoxicity data (OECD, 2003 (a)). Assessors should be aware of developments and debate in this area, and be ready to adopt new procedures when appropriate.

In the interim, as a general rule DEWHA will assume that a greater than 10% difference is biologically significant, even if this is not determined to be statistically significantly different (often due to the weak power in the statistics). This will be determined with some flexibility depending on the end point, eg symptoms of phytotoxicity versus biomass loss in plants toxicity testing.

6.1.4 COMMENT ON EMERGING ISSUES

Developments in effects assessment include enhanced understanding of particular mechanisms of action or toxicity, such as endocrine disruption, immunosuppression or immunological effects, and neurological effects. These mechanisms can have potential effects that are relevant to an assessment, however, not all of the endpoints measured in such tests are relevant for regulatory purposes. Regulatory authorities internationally are focusing on developing guidance material on the relevance and robustness of test methods and endpoints for these mechanisms, and how to best incorporate relevant details into data requirements.

For example, endocrine disruption caused by a chemical can be viewed as an effect. Consequently, it can be assessed within the normal framework outlined for effects assessment. However, the difficulty with such an assessment is the lack of test data available to adequately support assessment. Some information may be gleaned from studies received as part of the data package, for example, long-term studies on fish, aquatic invertebrates,

birds and some mammal studies. However, the extent to which these studies provide useful information is questionable (see below).

To address this problem the OECD established a Special Activity on Endocrine Disruptor Testing and Assessment in 1996. Test guidelines and documents related to this activity may be found at <http://www.oecd.org/document/62/0,2340,en_2649_34377_2348606_1_1_1_1,00.html> and assessors should be familiar with developments in this area. A useful monograph available from this site is the *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals* (OECD, 2002). The content of this document is summarised in Appendix IV.

A further useful source of information in the area of endocrine disruption is the *Global Assessment of the State-of-the-Science of Endocrine Disruptors* (Damstra *et al*, 2002).

Work on neurological and immunological effects assessment has proceeded more slowly, but developments by OECD in these areas should be monitored and acted upon by assessors as appropriate.

6.1.5 COMMENT ON FORMULATIONS

Often the active constituent of a product is the chemical most likely to cause a toxic effect. However, on occasion, solvents, adjuvants and other co-formulants may also have toxic effects. In some cases, they may be more toxic than the active constituents. Assessment of formulations is generally considered on a case-by-case basis, because the APVMA data requirements are silent on the issue of when formulation toxicity testing should be performed. As noted in MORAG:

> Toxicity information on the formulation to be used is also an important consideration, including for combination products to clarify whether the toxic effects exerted by the different active constituents are additive or not. Toxicity of the mixture **must** also be addressed in the case of deliberate tank mixes where the label instructs that for general, or for a particular use, the product and its active(s) must always be mixed with another different active constituent contained in existing products.

Expert judgement is used to determine if toxicity testing on the single active constituent formulation is required to supplement that on the active ingredient, because there are no formal criteria for such testing in Australia. Nevertheless, the EU offers guidance in this regard as follows (EC, 2002 (c)):

> Acute toxicity studies are not required for every formulation. However, co-formulants and solvents in formulations may significantly increase or decrease the acute toxicity of the active substance. There is some difficulty in predicting which types of formulations are critical in terms of such interactions. If the formulated product contains more than one active substance, this also complicates the prediction of toxicity using data on the individual active substances.
>
> Acute toxicity data on a *formulation* takes the toxicity of the co-formulants into account, as their toxicity will also be exerted in the tests. If the active substance is more acutely toxic when it is formulated, then risk quotients should be calculated on the basis of the data for the product.
>
> In principle, the following requirements should apply. Tests should be carried out on one species from each of the three groups of aquatic organisms (fish, aquatic invertebrates and algae). Where the available information on the active substance indicates that one group is clearly more sensitive, then tests on the most sensitive species of the relevant group should be carried out. In this context, the most sensitive group is defined as being at least 100 times more sensitive than the next most sensitive. If the least sensitive group is at least 100 times less sensitive than the most sensitive, then formulation data are not required on the least sensitive group. If the most sensitive species tested with the active substance is *Lemna*, *Chironomus* or other species, then these should be tested with the formulation. For poorly soluble chemicals, tests on the formulated product may be required for a group that does not show toxicity for the active substance at the solubility limit.
>
> If the formulated product contains two or more active substances, and the most sensitive taxonomic groups for the individual active substances are not the same, then formulation toxicity data are required on all three groups.
>
> There is some scope for extrapolation of toxicity data between similar formulations. In addition, in some cases it may be possible to reliably predict the toxicity of a simple formulation from data on the active substance and information on the co-formulants. The notifier should justify such approaches in reasoned cases.

6.2 TERRESTRIAL ECOTOXICITY DATA

Agricultural chemicals may pose hazards to non-target terrestrial plants and animals. In order to determine these hazards, applicants are required to provide environmental toxicology data on the effects of their pesticides on birds, other terrestrial vertebrates, invertebrates and plants (refer to Chapter 3), as mentioned briefly below.

Among birds, the bobwhite quail and mallard duck are typical test species. Acute and chronic oral and dietary toxicity tests, and reproduction tests (as required), are conducted with each of the two species. The reproduction test is designed to check for the mortality of adults and chicks (both matched and unhatched), as well as such sub-lethal effects as reduced egg production and thin eggshells.

For terrestrial vertebrates, the guidelines covering toxicity testing on birds can be complemented with those conducted on mammals. The main objective for mammalian toxicity tests is, obviously, the assessment of human health effects; however, the same test data can be used for covering terrestrial wild mammals if required (OECD, 2004 (b)). Toxicity tests for birds and mammals are discussed further in sub-sections 6.2.1 and 6.2.2, respectively.

In addition to data on birds and mammals, laboratory studies are also conducted (as required) to determine toxicity to:

- invertebrates, such as bees and other insect pollinators (Sub-section 6.2.3.2)
- predatory or parasitic insects and predatory mites (Sub-section 6.2.3.3)
- earthworms, which are important for soil fertility (Sub-section 6.2.3.4)
- soil micro-organisms (Sub-section 6.2.3.5)
- non-target terrestrial plants (Sub-section 6.2.4).

6.2.1 BIRDS

DEWHA may require the following acute oral, subacute dietary, and chronic tests in birds in order to determine likely short and long-term effects in non-target bird species:

Avian acute oral LD_{50} test is conducted with either an upland game bird (eg, bobwhite quail) or a waterfowl species (eg mallard duck). It is an acute, single-dose laboratory study that is designed to determine the amount (dose) of pesticide that will cause 50% mortality (LD50) in a test population of birds.

Regurgitation may significantly reduce the dose that the bird actually receives. Therefore, the occurrence of regurgitation or emesis should be assessed during the evaluation of avian acute oral tests. If it has, it may be appropriate to repeat the study using birds that do not regurgitate, in particular if a high-risk use – such as seed treatment – is being assessed (EC, 2002 (b)).

Avian dietary LC_{50} test is conducted with an upland game bird (eg bobwhite quail) and a waterfowl species (eg mallard duck). It is an acute, eight-day dietary laboratory study designed to determine the amount of pesticide in food that will cause 50% mortality (LC50) in a test population of birds.

Avian reproduction test uses both an upland game bird and a waterfowl species. It is a laboratory study (usually 20 weeks) designed to determine the amount of pesticide that will harm the reproductive capabilities of a test population of birds. Reproductive impairment is measured in terms of number of eggs laid per hen, number of cracked eggs, number of viable embryos, live three-week embryos of viable embryos, normal hatchlings of live three-week embryos, and number of 14-day-old survivors. This test is used to determine the NOAEC (no observed adverse effects concentration) or LOAEC (lowest observed adverse effects concentration) for the above parameters.

Low acute and dietary avian toxicity are not sufficient to indicate a low reproductive toxicity. A reproductive test should be supplied unless it can be demonstrated that exposure of birds does not occur during the breeding season. When all relevant species are considered, the effective breeding season could be rather long. Therefore, even short exposure periods may give rise to concern with regard to potential reproductive effects. Thus, in the case of foliar applications during the breeding season, the test should normally be required even if only one treatment per season is intended. Reproductive data should always be provided for chemicals that are persistent or have a bioaccumulative tendency (EC, 2002 (b)).

Simulated or actual field testing is conducted either to quantify the actual risks in the field or to show that risks in the field under actual use are different than in the laboratory. These are higher tier tests and are seldom required. Generally, the methods for higher tier tests (which may also include such studies as palatability/ avoidance tests and pen/cage tests) are not standardised. Consequently, expert judgement is required to assess study quality.

6.2.2 MAMMALS

This is an area of assessment not always considered in Australian assessments unless the use pattern assures mammalian exposure (eg rodenticides, grain treatments, etc.). Guidance in this area is taken from the US EPA <http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_eco.htm#HDEUED>, and the EU (EC, 2002 (a) and 2002 (b)). Generally, the following studies are used to evaluate effects on small mammals:

Mammalian acute test is an acute, single oral dose test conducted in laboratory rats and mice. This test measures the amount of pesticide in a single oral dose that will produce mortality in 50% of the test animals (LD50).

Mammalian sub-acute test measures sub-acute dietary effects in rats and mice. That is, it measures the amount of pesticide in the diet that will cause mortality in 50% of the test animals (LD50). In the EU, the short-term assessment is covered by the acute and long-term assessments.

Mammalian chronic tests measure reproduction and developmental effects in mice and rats. As described in the EU methodology, the NOEL is based on the most sensitive end-point of relevance for survival rate, reproduction rate and development of individuals. Examples are multi-generation or teratology studies on mammals. In each category the toxicity of the most sensitive test species is used.

With regard to the long-term scale in mammals the Tier 1 assessment is conducted with a dose level that represents the toxicological NOAEC. If the resulting risk quotient is above the corresponding trigger, then the ecological relevance of end-points should be re-evaluated.

6.2.3 NON-TARGET INVERTEBRATES AND SOIL MICRO-ORGANISMS

6.2.3.1 Introduction

Soil is an extremely heterogeneous compartment when compared to the aquatic compartment. Soils vary in the mineral and organic matter content and sizes of particles, aggregates, pores and micropores. In addition, they contain air and water. These create a wide spectrum of environmental conditions that influence chemical behaviour in a number of ways. Chemicals may be adsorbed to the organic matter or to mineral components of the soil, thus reducing their bioavailability. Soil pores and micropores may act as traps for desorbed matter. Substances may at the same time encounter aerobic and anaerobic conditions. Soil water may therefore be quite different in chemical reactivity from surface water. Thus, bioavailability, ageing and other phenomena influence and modify the effects of chemicals in soil (ECETOC, 2002). Consequently, the terrestrial effects of a substance are determined by a combination of its intrinsic toxicity and bioavailability. Experimental effects data (where exposure is through the soil) are always a combination of both factors.

In interpreting effects data, and possibly considering the need for higher testing based on these data, assessors consider a number of modifying factors. While intrinsic toxicity and bioavailability are the main drivers for soil hazard, there are a number of other parameters/properties that influence the hazard that a substance may pose to organisms in the soil. These properties include potential for long-term effects and elimination processes including degradation, volatilisation and leaching.

6.2.3.2 Bees

Acute toxicity to bees

If honeybees are likely to be exposed to the active substance then both acute oral and contact toxicity tests must be conducted, as the toxicity by one route of exposure cannot be predicted from the other. Where there is only one relevant route of exposure (eg oral exposure in the case of soil application), testing can be restricted to this exposure route. If there are problems with solubility of the active substance, then the test should be conducted with a representative formulation.

Bee brood feeding test

This type of testing is not commonly sought with Australian assessments but may be considered on a case-by-case basis. The test method of Oomen *et al*, 1992 is recommended in EU data requirements for insect growth regulators, and is a worst case screening test. If no effects are found in this test then it can be concluded that no brood damage will occur when using the product. However, if effects are found, then further cage/tent/tunnel or field studies are necessary to evaluate the risk under more realistic conditions. If toxicity to honeybee broods can already be predicted from the mode of action of the compound, testing may immediately start with cage/tent/tunnel or field trials.

Residue test

Aged residue tests may be valuable as an additional tool for risk assessment. However, no specific validated methods are available as yet. The test should be designed to assess the duration of effects due to residual traces of plant protection products on the crop.

Testing of systemic plant protection products

For soil-applied systemic plant protection products (eg plant protection products applied as seed dressing) the acute oral toxicity of the active substance(s) has to be determined. If potential risks to honeybees are identified (ie very low LD50) then realistic exposure conditions should be taken into account. That is, realistic exposure concentrations expected to be present in nectar and pollen, as indicated by residue studies. If a risk is indicated, then higher tier studies (cage/tent/tunnel or field studies) with realistic exposure scenarios should be performed.

6.2.3.3 Other non-target arthropods

The risk to other non-target arthropods is routinely assessed where pesticide use will result in exposure to such organisms. Standard tests received in this area include measurement of effects to organisms such as predatory mites (*Typhlodromus pyri*), parasitic wasps (*Aphidius rhopalosiphi*), rove beetles (*Aleochara bilineata*), carabid beetles (*Poecilus cupreus*), collembola (*Folsomia candida*) and Lycosid spiders (*Pardosa sp.*). Such tests tend to follow IOBC/BBA guidelines. No dose-response relationship can be established when these tests are limit tests (as is the case with soil micro-organisms and usually with beetles and spiders). Effects should be expressed in terms of loading rates, because measurements of concentration include bioavailable and non-bioavailable fractions and so do not reflect actual exposure (ECETOC, 2002).

International approaches

The US EPA does not usually require testing on beneficial arthropods such as predatory mites. Rather, they may require testing for effects in insect pollinators, such as honey bees, when the typical end-use product (TEP) is intended for outdoor use and honey bees may be exposed to the pesticide <http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_eco.htm#WSAN>.

Developments in the approach taken by the EU in assessing other arthropods are worth noting. The following discussion is based on two source documents, EC, 2002 (b) and Candolfi *et al*, 2001 – referred to below as ESCORT 2. These documents should be consulted for a fuller explanation.

The risk to non-target arthropods is routinely assessed in Europe. Data requirements state that data on two sensitive standard species as well as data on two crop relevant species are required. If effects are observed with species relevant to the proposed use, then further testing may be required. Where significant effects have been observed then the toxicity of the product to two additional species must be investigated. While EU guidance currently references the SETAC Guidance document on regulatory testing procedures for pesticides with non-target arthropods (ESCORT, Barrett *et al.* 1994) as a source of guidance for testing, several limitations with the SETAC guidance document have been identified. These can be summarised as follows:

- The objectives of the testing scheme are not clear, for example, it does not precisely discriminate between non-target arthropods in a general context and beneficial arthropods in an agricultural or IPM context.
- The trigger value for Tier 1 data (30% effects) leads to excessive higher tier testing.
- The single-dose laboratory data generated do not provide for determination of the intrinsic toxicity of the substance (except where is no effect and the test can be regarded as a limit test). In addition, this kind of testing is inflexible and does not allow a satisfactory risk assessment especially for off-field habitats.
- Uncertainty remains regarding data requirements, testing methodology and evaluation, especially for multiple application products (where currently life span, spraying interval and fate are ignored) and for off-crop habitats (where exposure scenarios and mitigation measures are not yet agreed).

Because of these limitations, a workshop, ESCORT 2, was held in 2000 that aimed to address these shortcomings. All EU member states along with representatives from industry and academia attended the workshop and revised the process by which the risk to non-target arthropods should be assessed. By building on the experience gained from assessing the risk to non-target arthropods, a new approach was proposed which offers a high level of protection, but is more focused and structured. This new approach is contained in a guidance document (Candolfi *et al.* 2001), which is hereafter referred to as ESCORT 2. Salient points are discussed below.

The process recommended by ESCORT 2 starts with glass-plate tests on the two standard sensitive species (*Aphidius rhopalosiphi* and *Typhlodromus pyri*). However, rather than a single rate study, a rate-response study is usually required. The end-point of these studies is LR50 values (ie values for the lethal rate that causes 50% mortality) which are compared to the predicted exposure both in-field and off-field. The assessment of risk for

arthropods living in- and off-field is conducted separately. If the resulting 'hazard quotient' (HQ) based on the standard tests is greater than or equal to 2 then further data and/or risk management measures are required. Note: The critical trigger of 2 was proposed on the basis of the available data. It was noted at the ESCORT 2 workshop that this value should be revised when suitable data are available. This document should be consulted for rationale behind the quotient value of 2.

It is proposed that for active substances and their associated product(s) being assessed, the risk to non-target arthropods both in and off-field should be adequately addressed. The guidance given below is in line with the recommendations of ESCORT 2.

Standard tests

Testing is always required where exposure of non-target arthropods is possible.

Standard tier 1 testing comprises glass-plate tests with *Aphidius rhopalosiphi* and *Typhlodromus pyri*. Preferably these tests should be designed as rate-response studies in order to determine the LR50 as this allows application of the data to different use scenarios and also to the risk assessment for off-crop areas. However, if the toxicity is expected to be low, then limit tests can be conducted at a rate equivalent to the maximum application rate multiplied by the multiple application factor (MAF). With substances suspected to have a special mode of action (eg IGRs, insect feeding inhibitors) tests should include sub lethal end-points and may also need other modifications.

Details on methods are given in the ESCORT 2 document.

Higher-tier tests

Higher-tier tests are required when a risk is indicated in lower assessment tiers. There are several options for higher-tier testing or combinations of adequate tests:

- extended laboratory tests (tests with natural substrate aiming at lethal and sub lethal effects)
- aged-residue studies
- semi-field tests
- field tests.

ESCORT 2 provides advice regarding the choice of studies and the selection and number of species. Usually these studies are conducted with one dose rate which matches the field application rate, taking into account multiple applications and the use of appropriate risk mitigation measures. Advice is given in ESCORT 2 regarding the appropriate rates to use in such studies.

In the case of extended laboratory tests, the dose rates may be greater than the field rate in some cases because of implementation of a correction factor (default value = 5). In this situation, it is suggested to test at the maximum rate including the multiple application calculation. In the case of extended laboratory studies a dose response design may be more informative than a one-dose design.

6.2.3.4 Earthworms and other soil macro-organisms

Acute effects on earthworms

Testing is always required where contamination of the soil is possible.

Sub-lethal effects on earthworms

The requirement for this test depends on the exposure pattern of the active substance ('continued or repeated exposure'). Australia does not have any formalised trigger values to ascertain the need for such tests; rather, the need is based on expert judgement. The EU (EC, 2002(b))has proposed the following triggers for persistence of the active substance and the number of applications as described below:

- the test is not required when the DT90 (time taken for 90% of a substance to be removed from a system) in field studies is less than 100 days, and the number of applications is less than three
- the test is always required if the DT90 in field studies is above 365 days (regardless of the number of applications)
- the test is always required if the number of applications is greater than six (regardless of persistence)
- if the DT90 in field studies is between 100 and 365 days and/or the number of applications is between three and six, a case-by-case decision is made.

The test is also required if the assessment of the acute risk gives a risk quotient >0.1 (refer Chapter 8).

The test should preferably be conducted as dose-response test. When planning the test, the upper concentration level must be chosen to be high enough in order to be able to judge whether the long-term risk quotient meets the trigger (see Chapter 8). Exposure under field conditions may be elevated due to repeated applications and toxicity figures may need to be corrected for the fraction of organic carbon (Foc). If available, data from field dissipation studies should be considered.

Earthworm field studies

Again, there are no formal criteria for requiring this type of test in Australia. Rather, the need is determined through expert judgement. In the EU, the need for this study is triggered where the long-term toxicity exposure ratio, TER, is <5 (chronic Q >0.2). However, in such cases it should be checked whether there are other options for refinement.

The study should reflect the intended use of the compound, the environmental conditions and species that will be exposed. For example, if the chemical is intended to be applied in the arable situation then in the study it should preferably be applied to bare soil as opposed to grassland where it may become bound to the surface thatch. Analysis of the soil would assist in confirming whether the field study is appropriate for the intended arable crop use. In terms of dosage, the test should include the highest exposure according to the intended use of the product. That means that multiple applications should be made where relevant, and crop interception should also be considered. In addition, a rate equivalent to the long-term plateau concentration should be added if accumulation in soil is expected. The type of application to be used in the field should be used in such studies (surface application, incorporation, etc.).

Other soil macro-organisms

Again, some guidance for the types of tests that may be useful in this area can be found in EC, (2002b) where there may be additional data required for soil organisms contributing to organic matter breakdown, depending on active substance degradation rate, and on available information with regard to effects to various organisms. The following guidance is offered:

> Principally the risk to this group of organisms, which include soil mesofauna and macrofauna, could be determined either at a species level or at a functional level. While a candidate test for the former would be a Collembola reproduction test or a test on gamasid soil mites, a candidate for the latter would be the "litter bag" test.
>
> Persistent active substances or persistent metabolites (DT90 >100 days) are of special concern as influences on organisms can continue to act over generations and may have multiple effects, and any recovery could take an unduly long time. Therefore, a higher degree of scrutiny is needed to assure that soil organisms are not affected.

The guidance document recommends the following procedure:

> *a) Collembola reproduction test or test on gamasid mites*
>
> Testing is required where contamination of soil is possible and DT90 is between 100 and 365 days and the standard hazard quotient for arthropods (*Typhlodromus* and *Aphidius*) >2. This test is used as a potential waiver for the litter-bag-test (see next point); so, if the litter-bag test is triggered anyway by other criteria (effect on soil micro-organisms >25 % or TER for earthworms <5) then this test could be omitted.
>
> A suitable protocol for the Collembola test is the ISO method 11267:1999.
>
> *b) Litter bag test under field conditions*
>
> Testing is always required where contamination of soil is possible and DT90 is > 365 days or mineralisation is <5% in conjunction with bound residue formation of >70%. Testing is conditional where DT90 is between 100 and 365 days; in such cases the following auxiliary criteria are applied:
>
> - Effects on soil microorganisms >25% after 100 d
> - or long-term TER for earthworm <5
> - or TER for Collembola or soil mites <5
> - Principally this means that in the intermediate persistence range a litter bag test is not required if the above mentioned groups of organisms pass the standard tier 1 assessment.

6.2.3.5 Soil micro-organisms

Micro-organisms play an important role in breakdown and transformation of organic matter in fertile soils, with many species contributing to different aspects of soil fertility. Any long-term interference with these biochemical processes could potentially interfere with nutrient cycling, which could in turn alter soil fertility. Transformation of carbon and nitrogen occurs in all fertile soils. Although the microbial communities responsible for these processes differ from soil to soil, the pathways of transformation are essentially the same (OECD TG 217).

Because of the importance of soil micro-organisms, testing on such organisms is required where use of a chemical will result in soil exposure. The two most likely tests to be available measure the carbon transformation activity (eg OECD TG 217) or the nitrogen transformation activity (eg OECD TG 216) of soil micro-organisms.

With these types of tests, evaluations of test results with agrochemicals are based on relatively small differences (ie average value ± 25%) between the carbon dioxide released or the oxygen consumed in (or by) control and treated soil samples or by nitrate concentration between control and treated samples. For this reason, large variations in the controls can lead to false results. Therefore, the variation between replicate control samples should be less than ± 15%.

An agrochemical product can be evaluated as having no long-term influence on carbon or nitrogen transformation in soils when tests indicate that the difference in the respiration rates or the rates of nitrate formation between the lower treatment (ie the maximum predicted concentration) and the control is equal to or less than 25% at any sampling time after day 28. In the case of tests with chemicals other than agrochemicals, the EC50, EC25 and/or EC10 values are used for evaluation.

6.2.4 NON-TARGET TERRESTRIAL PLANTS

This is an area of emerging science and methodology. As mentioned earlier (Section 5.4.3), this area has been the focus of concerted international efforts to determine appropriate test protocols, but much work remains to be done. Effects assessments methods should therefore be amended as work in this area proceeds.

The need to assess effects to non-target plants in Australian assessments has been driven by mode of action (eg herbicides are considered much more carefully than other groups of pesticides) and exposure arguments. The APVMA data requirements state that results of non-target vegetation toxicity data may be required, either as laboratory data or observations (generally from efficacy trials).

This has resulted in some variation in the methods for assessing effects to non-target plants in the past. Generally, the only available data in this regard are seedling emergence or vegetative vigour tests performed to US EPA guidelines or to OECD TG 208 (recently revised) and the newly adopted OECD TG 227. When these data are not available, the effects and corresponding risk characterisation assessment is usually performed qualitatively, possibly based on extrapolation from aquatic plant results.

The following Section describes current effects assessments for non-target terrestrial plants, based on international approaches.

6.2.4.1 International approaches

As mentioned earlier, Section 5.4.3 described an international workshop convened by the Office of Pesticide Programs (OPP) and the Office of Pollution Prevention and Toxics (OPPT) within the U.S. Environmental Protection Agency (EPA) on assessing the ecological risk of chemicals to non-target terrestrial plants. The conclusions of this workshop were as follows:

> It was concluded that, based upon incident data, current test protocols are not sufficiently predicting effects upon the development and reproduction of terrestrial plants, and that test methods to address these effects are needed. Tests with perennial terrestrial plants are needed. More than one measure of effect should be used in any toxicity test.
>
> Extrapolation from tested species to other species continues to be an issue. Evidence suggests that important effects are being missed when only using the currently required terrestrial species, but more research is needed. Variation in sensitivity between different species will also be dependent upon the mode of action of the pesticide. The need for a variety of species versus a few good indicator species is likely to vary depending upon the problem formulation. To select additional species, both phylogenetic diversity and relevant ecology may need to be considered. Thus, it would be useful to develop a decision rationale for selecting species that represent both the ecology and taxonomy at risk.
>
> The form of the chemical that should be tested (technical active constituent or formulated product) should be driven by the problem formulation. Aquatic species are typically exposed to the active constituent while terrestrial plants are typically exposed to the formulated product. In the absence of evidence that various formulations have the same toxicity, data needs to be generated. It may be appropriate to develop data for the active constituent and then develop bridging data for formulation(s). Testing of isomers, degradates, and co-

formulations is currently requested on a case-by-case basis. The chemical and fate properties of the pesticide should be used to direct the exposure and effects assessments. These include solubility, volatility, adsorption, degradation, etc. It was noted that the exposure information is generated based upon the technical active constituent, not the formulated product.

Research is needed to develop new phytotoxicity test methods, validate existing methods, and generate test results. Some of the most important research needs were: tests for terrestrial plant development and reproduction; field tests, including multi-species approaches; monitoring tools; improved incident reporting; research on alternative test species, on relative sensitivity, and on approaches to selecting test species; methods to evaluate recovery; and greenhouse-to-field extrapolations.

Incident data can be used to support or modify the conclusions of the risk assessment, to help identify effects that were previously missed, to direct the risk assessment for chemicals similar to existing chemicals, or to focus research. Problems with incident data include the lack of reporting, lack of details in the reports, and the difficulty in establishing cause and effect. Better investigation, more interaction between agencies, and improved reporting is recommended.

Uncertainty is addressed differently in different tiers of the risk assessment. Uncertainty factors have been used to extrapolate between species when available data are limited. The use of uncertainty factors can be reduced when there are sufficient data to use the species sensitivity distribution. There was no agreement on the number of species that would provide a reasonable distribution of sensitivity. The confidence in the effects data must be compared to the confidence in the exposure data, and a decision made as to where it is most crucial to reduce uncertainty.

Such conclusions highlight the ongoing need for refinements of methodology. Nevertheless, there are existing methods for assessing effects on non-target terrestrial plants used by the US EPA and the EU that can be drawn upon.

The following discussion on effects assessment to non-target terrestrial plants is based on current methodology used in the US <http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_eco.htm> and the EU (EC, 2002 (b)).

In the USA, the plant-testing scheme is tiered, such that a limited number of species are tested in the first level. If the first level tests show effects, then additional plant species are studied at a higher level. In these tests, multiple species of herbaceous plants (crop species) are tested for seedling emergence and vegetative vigour (NOAEC/LOAEC or EC25).

The EU does not have specific data requirements for non-target plants. However, there is a need to report all potentially adverse effects and to undertake additional studies where there are indications of such effects. Therefore, a tiered approach is suggested starting with available data and proceeding to further steps when necessary. Data are not required where exposure is negligible, for example, in the case of rodenticides, substances used for wound protection or seed treatment, or in the case of substances used in stored products or in glasshouses.

Tier I

- **US EPA:** Tier I testing is a greenhouse or growth chamber test that consists of two parts: a test for seedling emergence and a test for vegetative vigour. Seedling emergence is a 14-21 day test, examining the effect concentration that affects 25% of the exposed plants (EC25) for percent emergence, plant height, plant dry weight, percent visual phytotoxicity, EC05, and NOAEC (No Observed Adverse Effects Concentration). Vegetative vigour is a 14-28 day foliar spray test that is used to determine the EC25 for plant height, plant dry weight, and percent visual phytotoxicity, EC05, and NOAEC. The tests are usually conducted with crop species, such as corn, soybeans, and a root crop. Seven other species that may be used include tomato, cucumber, lettuce, cabbage, oat, ryegrass, and onion.
- **EU:** For the first tier, a preliminary assessment is conducted using available information. Preference is given to screening data; there should be at least six species from different taxa tested at the highest nominal application rate (1 x). These data could be supplemented by further information on efficacy, selectivity, phytotoxicity, etc. included in the biological dossier or obtained from the different field assays such as efficacy trials, residue studies, environmental fate and ecotoxicological studies, etc. Herbicides and plant growth regulators will require Tier II testing so Tier I testing is usually not considered necessary.

Tier II

- **US EPA**: Tier II testing is a greenhouse, growth chamber, or small plot test that consists of three parts: a seed germination test, a seedling emergence test, and a vegetative vigour test. These tests evaluate the effects of multiple dosage levels on plant growth.
- **EU**: If a potential risk is identified (ie more than 50% effect for one or more species at the maximum application rate), then specific information on the toxicity of the substance to terrestrial plants should be

requested. The second tier considers laboratory assays on a selection of plant species. It is recommended that dose-response tests be conducted on 6-10 plant species that represent as many taxonomic groups as possible. In order to generate data that are useful for probabilistic approaches, there should not be a focus exclusively on species assumed to be the most sensitive. If, from the screening data, a specific mode of action is evident, or strong differences in the species sensitivities are identified, then this evidence should be used in the selection of the appropriate test species. This may be especially true if non-herbicides reach Tier II testing.

For foliar applications, the bioassays should be conducted by spraying the product on the plants. Bioassays should reproduce, as far as possible, realistic exposure conditions including spray drift. Soil application should be chosen if that is more appropriate with regard to the use pattern. The test substance should be the lead formulation (or another formulation) because formulations contain, all those components and co-adjuvants required for maximising biological activity in addition to the active constituent. For systemic products applied on the ground/soil, the tests should reproduce this application pattern.

Suitable test methods are the new draft OECD Guideline 208 and the OPPTS guidelines of the US EPA.

Tier III

- **US EPA:** Tier III testing includes terrestrial field tests that may be required on a case-by-case basis if terrestrial plants show greater than 25% adverse effects on plant growth. These tests provide critical information on harmful effects to plants during all stages of development.
- **EU:** The third tier requires semi-field or field assays, to study the effects observed on non-target plants during realistic applications. Such studies are time-consuming and expensive; before undertaking them it should be checked whether there are options for the refinement of exposure and/or effects. Furthermore, as for all other non-target organisms, field or semi-field studies are not required if the risk indicated in the Tier II assessment could be managed by risk mitigation measures.

  Field or semi-field studies with non-target plants are not standardised. Consequently, protocols should be agreed between the company and the assessing agency. Generally, effects on plant abundance and biomass production at different distances from the crop or at exposure levels representing different distances from the crop should be analysed. These studies are compatible with most semi-field and field studies.

6.3 AQUATIC ECOTOXICITY DATA

6.3.1 INTRODUCTION

As noted at the start of this chapter, the assessment of environmental effects involves evaluating data on lethal and sub-lethal effects in acute and chronic toxicity laboratory tests. As also discussed in the preceding section on terrestrial effects, in some cases field data may also be required. For assessing effects on aquatic organisms, essential data for aquatic ecotoxicity cover the three trophic levels: primary producer (algae/aquatic plants); primary consumer (aquatic invertebrates); and secondary consumer (fish). Standard tests for these three trophic levels are discussed below. The need for higher tier testing is dictated by acute test results, use pattern and properties of the chemical itself.

For primary producers, results of toxicity tests on freshwater and marine algae and aquatic vascular plants are evaluated. The requirements for the number of species tested are dictated by the chemical type and use pattern (eg herbicides with inland and coastal use should provide data on freshwater and saltwater algae/aquatic plants).

Information on primary consumers, such as acute and chronic toxicity to aquatic invertebrates, (eg freshwater – *Daphnia* sp.; saltwater – mysid shrimp) is routinely provided. Additionally, test data on molluscs (shellfish) should be provided and evaluated for pesticide uses involving marine deposition.

Acute and chronic-toxicity tests on secondary consumers are conducted with both cold- and warm-water fish species (generally, rainbow trout and bluegill sunfish, respectively). Data on toxicity to salt-water species may also be available and should be provided when the proposed use pattern will result in marine exposure.

Several concepts should be considered prior to the assessment of effects for these three trophic levels, which will be discussed in Sub-section 6.3.4. As we have already noted, the first step of the effects evaluation process should be an evaluation of available ecotoxicity data for adequacy and completeness. Evaluation of data has been discussed in Chapter 4. However, further to this, modifying factors that are important in aquatic hazard classification - such as partition coefficient (log Kow), biodegradation and bioaccumulation - should be taken into consideration (Sub-section 6.2.2). Comment on use of QSARs follows in Sub-section 6.2.3.

6.3.2 EVALUATION OF DATA USED FOR THE ASSESSMENT

Before conducting an effects assessment, data should be evaluated for their adequacy (see Chapter 4). Specific considerations for data evaluation, described in OECD, 2004 (a), are summarised below.

For the interpretation of data, the key aspects of the study methods that affect study quality should be examined, such as measured or nominal concentration, control response, use of insensitive species, and water quality values. End-points that have direct ecological relevance (eg survival, growth, reproduction) should be given more weight than other end-points (eg biochemical parameters). Consideration of test species is also important: for example, chronic studies should be conducted with the most sensitive species in the acute tests.

If multiple data are available for the same species, the following procedure is proposed:

- if these data are based on the same effect parameter (end-point) and the same time period, the geometric mean value can be used
- if different effect parameters or different exposure times are used, only the lowest value from the longest test time should be used taking into account the importance of the end-points and the exposure periods in the various tests.

Octanol-water partition coefficient (K_{ow}): The octanol-water partition coefficient is an important parameter in initial hazard assessment, and therefore should be examined carefully. For example, determination of Kow by the shake flask method is not suitable for highly hydrophobic chemicals (log Kow > 5). For those chemicals, the slow stirring method or generator column method can be used. Testing using HPLC (OECD TG 117) is also appropriate. It should also be noted that test methods for determining log Kow may not be appropriate for surfactants, polymers, inorganics, and organometallics.

Bioaccumulation: Bioaccumulation occurs through multiple routes of exposure including uptake of food and sediment/soil. For most organic substances, uptake from water (bioconcentration) is believed to be the predominant route of exposure, although more recent work also considers bioaccumulation in air breathing animals (see Appendix V). Data on bioconcentration can be obtained by experiment or through a QSAR equation using Kow. It should be noted that simple bioconcentration QSARs often cannot predict the bioconcentration factor (BCF) of extremely hydrophobic chemicals under field conditions. If more than one BCF is available for the same species, the geometric mean for the species could be used; however, the test concentration used in determining the BCF should be taken into account. BCF values are more often available for fish, although results may also be available for other species (blue mussel, oyster, and scallop).

Further guidance on the interpretation of bioaccumulation data can be found in OECD 2001(a) and 2001(b). The GHS also provides a discussion on bioaccumulation that will assist assessors. This is reproduced in Appendix V. As well, chronic toxicity tests are preferred for persistent or bioaccumulative chemicals. For some of these chemicals, the standard 96-hour exposure period of acute tests may not be of sufficient duration to adequately determine their effects.

Water solubility: The water solubility of the test substance must be measured or predicted. The effect concentration derived from the test must not significantly exceed the solubility limit. Test results using solvents should be treated with care. For further guidance on testing of difficult substances, see OECD, 2000.

6.3.3 USE OF QSAR

There is a strong preference for using measured data in effects assessment. However, when data are insufficient (such as no data, data for only one test species, or when the measured data for a species is deemed to be unacceptable) then estimation using QSARs may be used. Such an approach is seldom used for assessment of agricultural chemicals. Chapter 3 provides general guidance on use of QSARs.

In regards to predicting toxicity, QSARs can be applied to chemicals with a common mode of toxic action, such as narcosis where the mechanism is dependent on a chemical's hydrophobicity (eg, log Kow). However, agricultural pesticides or veterinary medicines are often biologically active with more specific modes of action, making the use of a QSAR approach less applicable. Again, experience shows this is not a common approach with agricultural chemicals and consequently is not dealt with in detail in this manual.

Specific considerations for QSAR use in estimating aquatic toxicity are described in some detail in the Environmental Risk Assessment Guidance Manual for Industrial Chemicals, so is not repeated here.

6.3.4 AQUATIC EFFECTS ASSESSMENT

The testing requirements for aquatic organisms are tiered and are based on the use and toxicity of the pesticide active constituent. Depending on the results of studies conducted at a lower level, testing can progress from basic laboratory tests to a progressively higher, more real-world situation. The following discussion on aquatic organisms effects assessment is based on current methodology in Australia and draws on guidance from the US EPA <http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_eco.htm#Ecotox> and the EU (EC, 2002 (c)).

6.3.4.1 Algae/aquatic plants

As discussed in Section 5.4.2, assessment of risks to non-target plants is an area of emerging science and methodology. While toxicity to algae/aquatic plants is a data requirement, the number of species to be tested and the need to consider other aquatic plants remains an area of expert judgement. As noted in MORAG, ecotoxicity tests for at least 4 algal species and *Lemna spp.* should be included for all herbicides and fungicides due to the potential for harm to these species from these types of pesticides. However, fewer data points may be adequate for other types of pesticides.

Algal tests are short term (3 to 4 days), but they provide both acute and chronic endpoints as growth during this period remains in the exponential phase. The preferred observational endpoint is the inhibition to the algal growth rate (ErC50) because it is not dependent on the test design, whereas biomass (EbC50) depends both on the growth rate of the test species as well as test duration and other elements of test design. Thus, the former is preferably used by DEWHA in its risk assessment.

Section 5.4.2 described an international workshop convened by U.S. EPA on assessing the ecological risk of chemicals to non-target plants. The conclusions of this workshop with regard to effects assessment of non-target aquatic plants were as follows:

> It was agreed that current testing protocols, which use algae and duckweed, are not sufficient to address impacts to higher aquatic plants. There is no evidence to suggest that terrestrial plants can adequately serve as surrogates for aquatic plants. Thus, there is a need for testing with rooted aquatic macrophytes. If there is a potential for marine exposure, the effects assessment should include more than the one species of marine diatom that is currently tested. Mesocosm and microcosm studies can also be useful.
>
> Extrapolation from tested species to other species continues to be an issue. There is evidence to suggest that important effects are being missed when only the currently required aquatic species are considered, but more research is needed. Variation in sensitivity between different species will also be dependent upon the mode of action of the pesticide. The need for a variety of species versus a few good indicator species is likely to vary depending upon the problem formulation. To select additional species, both phylogenetic diversity and relevant ecology may need to be considered. Thus, it would be useful to develop a decision rationale for selecting species that represent both the ecology and taxonomy at risk.
>
> The form of the chemical that should be tested (technical active constituent or formulated product) should be driven by the problem formulation. Aquatic species are typically exposed to the active constituent, while terrestrial plants are typically exposed to the formulated product. Testing of isomers, degradates, and co-formulations is currently requested on a case-by-case basis. The chemical and fate properties of the pesticide should be used to direct the exposure and effects assessments. These include solubility, volatility, adsorption, degradation, etc.
>
> Research is needed to develop new phytotoxicity test methods, validate existing methods, and generate test results. Some of the most important research needs were: tests for aquatic macrophytes; field tests, including multi-species approaches; monitoring tools; improved incident reporting; research on alternative test species, on relative sensitivity, and on approaches to selecting test species; methods to evaluate recovery; and greenhouse-to-field extrapolations.
>
> Incident data can be used to support or modify the conclusions of the risk assessment, to help identify effects that were previously missed, to direct the risk assessment for chemicals similar to existing chemicals, or to focus research. Problems with incident data include the lack of reporting, lack of details in the reports, and the difficulty in establishing cause and effect.
>
> Uncertainty was also discussed during the workshop. Uncertainty factors have been used to extrapolate between species when available data are limited. The use of uncertainty factors can be reduced when there are sufficient data to use the species sensitivity distribution. However, there was no agreement on the number of species that would provide a reasonable distribution of sensitivity. It was concluded that the confidence in the effects data must be compared to the confidence in the exposure data, and a decision made as to where it is most crucial to reduce uncertainty.

The following sub-section describes current effects assessments for non-target aquatic plants, including the current approaches used by the US EPA and the EU.

International approaches

The following discussion on current data required internationally is taken from US EPA guidance <http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_eco.htm#Ecotox> and EU Guidance (EC, 2002 (c)).

US EPA. In the USA there are currently five aquatic plants, including algae, for which testing is required, based on a tiered approach. This means a limited number of species are tested in the first level. If the first level tests show effects, then additional plant species are studied at a higher level. In these tests, multiple species of aquatic plants (algae and duckweed) are tested for effects on growth (EC50).

Tier I

Tier I tests are laboratory tests that evaluate the acute toxicity of pesticides other than herbicides at the highest application rate to a freshwater green alga (*Pseudokirchneria subcapitata*) and an aquatic macrophyte (*Lemna gibba*). For herbicides, five species are usually tested at the highest application rate: *Skeletonema costatum, Lemna gibba, Anabaena flos-aquae, Pseudokirchneria subcapitata*, and a freshwater diatom, usually *Navicula* sp.

Tier II

Tier II tests are dose- response tests that are designed to evaluate the acute toxicity of pesticides to five aquatic species: *Pseudokirchneria subcapitata* (a freshwater green alga), *Lemna gibba* (an aquatic macrophyte), *Anabaena flos-aquae* (a blue-green alga), *Skeletonema costatum* (a marine diatom), and an unspecified freshwater diatom, usually *Navicula* spp.

European Union: In the EU, a test with green algae is required in all cases. For herbicides, an additional test (conducted in accordance with internationally recognised guidelines) is required on a further algal species from a different taxonomic group. The second species should be from a group other than green algae, such as diatoms or the blue-green algae. Plant growth regulators should be treated in the same way as herbicides because they act on primary producers.

Both biomass (cell number) and growth rate should be reported. Biomass is often the most sensitive end-point of the two. As there is no clear evidence available to indicate which will be the most relevant end-point for the field situation, whichever figure is lowest should be used in the risk assessment. Toxicity values should be based on the period of exponential growth.

Tests on higher aquatic plants (macrophytes) have to be performed for herbicides. Tests should be conducted with *Lemna* sp. There is a suitable ASTM guideline, an OECD-guideline and an EPA guideline (draft OPPTS 850.4400) available for this purpose.

The number of fronds is the most important end-point in the *Lemna* test. However, if toxicity values for biomass or other end-points are lower then these may be used for the risk assessment, if appropriate. Where a high risk to aquatic plants is identified on the basis of the standard *Lemna* test, the notifier should consider providing further information to demonstrate that the risk to higher aquatic plants is acceptable. It may be possible to obtain information on the mode of action, the importance of the different routes of exposure and the range of sensitivity from effects seen in terrestrial plant tests.

Additional studies using a range of aquatic plant species may be required for highly active compounds. Where the justification for an acceptable risk is based solely on a *Lemna* recovery study, the relevance to other aquatic plants which do not have the same capacity for rapid reproduction and/or for which the sediment route of exposure may be important, must be fully addressed.

If there is evidence from efficacy data or data on terrestrial plants that the data for *Lemna* are not representative for other aquatic plant species (eg auxin simulators which can be more toxic to submerged plants than for *Lemna*) additional data with other aquatic plant species may be required on a case-by-case basis. The test protocol for such studies should be discussed the regulator/assessing agency because no internationally accepted guideline is available.

At present, laboratory toxicity methods with aquatic macrophytes taxa other than *Lemna* are at an early stage of development. A protocol using *Myriophyllum* is being developed. Such methods will require further research before it is possible to develop a harmonised guideline.

6.3.4.2 Aquatic Invertebrates

Acute toxicity tests

Acute toxicity tests to aquatic invertebrates are considered a fundamental data requirement. Tests use a freshwater invertebrate (*Daphnia* sp.) in an acute, 48-hour laboratory study to determine the concentration of pesticide in water that causes 50% mortality or immobilisation (EC50) in a test population of invertebrates. Where marine exposure is likely, saltwater invertebrates such as mysid shrimp and molluscs (eg eastern oyster)

are also required. Mollusc tests determine EC50 values usually based on the concentration likely to result in reduced shell growth.

Chronic toxicity tests

Chronic testing may be considered as Tier II and may include a *Daphnia* reproduction test.

Other considerations

EC, 2002 (c) discusses a range of issues relevant to assessors and assessors should make themselves familiar with this material. This document should be consulted for further guidance. Salient points are reproduced below.

In the EU, there is a requirement for studies on gastropod molluscs and insects if continued or repeated exposure is likely to occur. However, in general this requirement is limited to chronic tests. Acute tests with gastropods and insects are only required if direct uses in waterbodies are intended.

Accepted international guideline for a chronic test on gastropods is not currently available. Furthermore, gastropod molluscs are generally significantly less sensitive than *Daphnia*. Consequently, for uses where a direct application is made to water, the notifier may make a reasoned case as to why gastropod mollusc data should not be required. This could include acute toxicity data demonstrating the relative sensitivity of molluscs and *Daphnia* to the active substance. A chronic study should only be required if continued or repeated exposure is to be expected.

In the EU, *Daphnia* acute and chronic toxicity data (with their associated uncertainty factors) are suitably representative for aquatic insects and other invertebrates in the case of herbicides and fungicides. For insecticides it should be carefully considered whether additional data on aquatic insects should be required. Whilst *Daphnia* have been demonstrated to be representative for most insecticides, the toxicity of certain recent chemistries which have very specific, receptor-mediated modes of action (eg neo-nicotinoids) may not be well-represented by *Daphnia*.

Information on the mode of action of insecticides (from efficacy and non-target arthropod data) should be considered before deciding whether testing on an insect species is required. If the toxicity of an insecticide to *Daphnia* is low (48 h EC_{50} >1 mg/L, 21 d NOEC > 0.1 mg/L), this may indicate that *Daphnia*.are insensitive to that chemical. In such cases, an acute toxicity test should then be carried out with first instar (2-3 d old) *Chironomus riparius* (48 h water-only study). There is currently no guideline for such a study available, but in principle the tests should be conducted using similar methodologies as for *Daphnia*. The toxicity data from the most sensitive organism (*Daphnia* or *Chironomus sp*) should be used in the standard risk assessment for invertebrates. The usual triggers for further assessment would also apply. If a long-term/chronic study on insects is already available there is no need to require an acute one.

If the 48 h EC_{50} for *Chironomus sp* is at least ten times lower than the *Daphnia* 48 h EC_{50}, then a chronic study should also be conducted with *Chironomus sp* (see sediment organisms below). In these cases, the same triggers that are applied to *Daphnia* should be applied to the *Chironomus sp* data (ie risk quotients of 0.01 for acute toxicity and 0.1 for chronic toxicity in the EU, to account for further potential differences in inter-species sensitivity of insects).

For insecticides that are insect growth regulators (eg benzoyl ureas and similar classes), special consideration should also be given to the potential for effects on aquatic insects. Such compounds tend to have more pronounced effects over longer time periods than standard acute studies (due to their effect on moulting). Therefore, chronic studies with *Chironomus sp* should generally be conducted, unless it can be clearly demonstrated that the onset of effects is rapid and that *Daphnia* are of similar sensitivity to chironomids.

Finally, any data available on estuarine/marine invertebrates (eg *Mysidopsis bahia*, oyster embryo larval studies) should also be provided.

6.3.4.3 Sediment organisms

Requirements for sediment toxicity tests generally rely on expert judgement with the need for studies in this area being particularly important for insoluble persistent pesticides. These tests may consist of a one-part test of 10 or 28 days plus a three-part test (14 days for each part) for freshwater organisms, plus a one-part test of 10-28 days for estuarine/marine organisms. *Hyallela azteca* and *Chironomus tentans* are commonly used in the one-part test, and the end-points of interest may include survival, growth and emergence. The three-part test consists of a 14-day aqueous exposure with minimal sediment, a 14-day sediment exposure with spiked sediment, and a 14-day interstitial exposure with the compound added to overlying water. *Chironomus tentans* is commonly used in the three-part test, and effect levels may include LC50, EC50, and BCF. The acute sediment test for estuarine/marine organisms is usually conducted with amphipods and measures LC50, EC50, NOAEC, and LOAEC.

In determing the need for sediment toxicity data, the route of exposure is an important factor. Where exposure is primarly through chemical bound to soil/sediment (for example, run-off in the sorbed state), data based on OECD TG 218 are more appropriate as the test is performed with the substance pre-mixed with the test sediments. However, in the case of exposure directly to the water column (for example through spray drift), data based on OECD TG 219 should be generated.

Where sediment data are lacking, an initial screening level assessment may be made by estimating toxicity through equilibrium partitioning (EqP). Aquatic toxicity data are converted into values for bulk sediment based on the partitioning characteristics of the chemical. This method has been described in the corresponding industrial chemicals guidance assessment manual. In using it, assessors should bear the following in mind.

The EqP method can be applied to all chemicals (including toxic metals) for which water quality standards (also known as water quality criteria) have been derived. In this case, the water quality standard will be the corresponding aquatic LC/EC50 value (the lowest value available should be used). The method is applicable to marine and freshwater sediments and across sites. However, the following qualifications apply (EC, 2003 (a)):

- The formula only considers uptake via the water phase. However, uptake may also occur via other exposure pathways such as ingestion of sediment and direct contact. This may be important, especially for adsorbing chemicals, such as those with a log Kow >3. For these compounds the total uptake may be underestimated.
- There is evidence from studies in soil that the proportion of the total dose remains low for chemicals with a log Kow even up to 5. Although it is recognised that results for the soil compartment may not be extrapolated to the sediment compartment in principle, it is considered that the possible underestimation of exposure is acceptable when using the EqP method for chemicals with a log Kow of 3-5.
- For compounds with a log Kow >5 or with a corresponding adsorption or binding behaviour (eg ionisable substances), the EqP method needs to be used in a modified way.

In order to take uptake via ingestion of sediment into account, the PEC/PNEC ratio is increased by a factor of 10. This approach is considered only as a screen for assessing the level of risk to sediment dwelling organisms. If a ratio >1 is derived with this method, then tests with benthic organisms using spiked sediment may need to be conducted to support a refined risk assessment for the sediment compartment.

The effects concentration determined using the EqP method must be compared to expected whole sediment concentrations in order to determine the need for actual sediment testing. Australia considers the need for sediment toxicity tests on a case by case basis. The EU proposes some triggers as follows (EC, 2002 (c)):

- Triggers for sediment studies should take into account the potential for exposure via the sediment as well as potential for toxicity. For example, results of a water/sediment fate study (eg OECD 308) would require a test on sediment-dwelling organisms if >10% of applied radioactivity represented by the parent compound is present in the sediment at or after day 14, and triggers to identify potential risks to invertebrates for toxicity are met.
- For compounds which do not reach the 10% trigger but are applied more than once during the season, due consideration should be given to the potential for accumulation of residues in the sediment. Exposure triggers based on the water/sediment study are more difficult to apply to such use patterns because in the water/sediment study, typically only a single application is made. However, the EU may use the FOCUS Step 2 calculator which, as well as including drift, considers potential inputs from the soil compartment (via drainage/run-off). FOCUS stands for: *Forum for Co-ordination of pesticide fate models and their Use*, and details of the scenarios and models can be found at http://viso.jrc.it/focus/. The compound is partitioned between 30 cm depth of water and 5 cm depth of sediment, and is degraded. At Step 2, it is assumed that both the soil and water compartments experience no dilution, and that equilibrium develops between the sediment and water compartments, with concentrations only influenced by degradation.
- The FOCUS calculation partitions the compound between water and sediment and assumes that equilibrium exists (this would be worst-case because in nature dilution would be expected). Therefore the concentration in the water phase will reflect the 'bioavailable' concentration in the sediment (the EqP approach). Consequently, using the appropriate water phase concentration, *Daphnia* toxicity data, it is possible to determine whether there is potential for sediment toxicity. Hence, if the TERs (based on the maximum exposure concentration at Step 2 from the 'Step1_2 in FOCUS' calculator) for *Daphnia* are less than 100 or 10 for acute or chronic end-points respectively (equates to risk quotients of greater than 0.01 and 0.1), then testing of sediment dwelling organisms should be required.
- There are guidelines to prevent unnecessary testing with substances of low toxicity to invertebrates. For example, the NOEC in the chronic *Daphnia* test (or in a comparable study with insects when this group of

organisms is more sensitive) must be <0.1 mg/L before testing on sediment dwelling organisms is considered warranted. This number was chosen based on data from monitoring studies which indicated that it is unlikely that higher concentrations will often occur in surface waters. A recent review that compared toxicity data for *Daphnia* with that for sediment-dwellers supports the aforementioned approach (Streloke *et al*, 2002).

- For persistent substances, it may be justified to require a life-cycle test on chironomids in order to generate data on reproduction effects. However, a standardised test method is not available in EU guidance, and there have only been a limited number of studies published in the literature.
- For insecticides where it is possible that *Daphnia* are not a representative test organism, then acute toxicity data for *Chironomus riparius* can also be used to trigger long-term sediment studies. If the sediment exposure triggers are met and the TER resulting from the maximum PEC at Step 2 and the *C. riparius* 48 h LC50 is less than 100 (risk quotient >0.01), then long-term sediment testing is required.

The US EPA considers chronic sediment toxicity testing may be conducted when mortality exceeds 20% for any concentration level used in acute sediment testing. The chronic freshwater sediment test is a 65-day test that measures survival, growth rate, reproduction, and behaviour in *Chironomus tentans*. The estuarine/marine chronic test is a 28-day test that measures growth rate, reproduction, percentage of neonates that survive, and behaviour in *Leptocheirus plumulosus*.

6.3.5 FISH

Acute fish studies

Acute fish studies are a fundamental data requirement. The APVMA data requirements state these are necessary for both freshwater and marine species, although the actual species are not specified. Generally, freshwater tests are conducted on both a cold water (ie rainbow trout) and warm water (ie bluegill) species. Sheepshead minnow is usually used as a marine test fish. Such tests are acute, 96- hour laboratory studies designed to determine the concentration in water required to cause 50% mortality (LC50) in a test population of fish.

Chronic fish studies

Chronic testing may be considered as Tier II and may include a fish early life-stage (ELS) test and/or fish life-cycle (FLC) test. The ELS test is designed to determine the amount of pesticide that will adversely affect the reproductive capabilities of a test population of fish. Impaired reproduction and development in fish are measured in terms of number of embryos hatched, time to hatch, mortality of embryos, time to swim-up, and growth-weight and length.

The FLC test is designed to determine the amount of pesticide that will adversely affect the reproductive capabilities and various life stages of a test population of fish. Impaired reproduction and development are measured in terms of the number of days to complete hatching, number of embryos hatched, number of surviving larvae hatched, number of abnormal fish, length of survivors, weight of survivors, mean length and weight of first- and second- generation juveniles, number of survivors, and number of embryos.

The EU notes concerns regarding another possible test, the fish extended mortality test using OECD TG 204 (EC, 2002 (c)). This test has several limitations. Mortality is the only end-point covered, and the exposure duration may only be 14 days. Furthermore, the developmental stage tested is not particularly sensitive. However, in recent years, studies that have been conducted in accordance with this guideline have usually included exposure for 21 days, with mortality, growth and behaviour as end-points. In addition, the developmental stage of the rainbow trout to be tested is the same as recommended in the OECD guideline 215 for the 'Juvenile Growth Test', given that the weight of a 5 cm long rainbow trout (OECD 204) is in the range of 1-5 g (OECD 215). A combination of both guidelines is therefore considered most appropriate. Hence, the study should have 28-day exposure duration and include survival, growth and behaviour as end-points. In order to avoid unjustified animal testing, existing valid studies conducted in accordance with OECD 204 but lasting only 21 days can also be used to fulfill the data requirement.

It should be noted, that longer-term exposure might lead to sub lethal effects that are not covered by acute toxicity testing. Chronic toxicity tests are important as they provide valuable information on sub-lethal effects.

Generally chronic toxicity data must be provided unless it can be demonstrated that continued or repeated exposure is unlikely to occur, and the chemical is not deemed persistent. Australia does not have formal criteria for deciding whether chronic data are required; rather this is based on expert judgement.

Some more formalised criteria for chronic testing are used in the EU, and these are described below (EC, 2003 (c)): Using these criteria, a long-term/chronic test should be required if:

- The DT50 from the water-sediment study for the concentration of parent compound in the water column is ≥2days at an environmentally relevant pH in the range of 6-9. In practice, this means that chronic data are nearly always required.
- If the proposed use of an active substance involves more than one application per season, then long-term/chronic toxicity data are required, unless the DT50 in the water phase is <2 days and it can be clearly demonstrated that prolonged exposure will not occur due to the length of the spraying interval (this is the same as the US EPA requirement). Where such conditions apply, the potential risk from repeated acute exposures should be addressed in the assessment report on a case-by-case basis.
- For some active substances, the submission of the base set of data might not be sufficient to fully address the need for chronic toxicity data in order to complete the risk assessment. In these cases, the need for an ELS test or an FLC test should be considered. The ELS test is triggered if the acute LC50 for the active substance is less than the trigger value of 0.1 mg/L. This trigger should also apply to the FLC test.
- FLC tests may be required where the BCF is >1000, the elimination during the 14 day depuration phase in the bioconcentration study is <95% or the substance is stable in water or sediment (DT90 >100 days). However, taking into account that this type of study is difficult to conduct and often the results do not differ significantly from the ELS test, the FLC test may not be required if only one or two out of the toxicity, bio-accumulation and persistence triggers are breached. If all three triggers are breached then the test should be required. If effects on reproduction or the endocrine system could be anticipated (eg based on data from mammalian toxicology studies), the need for an FLC test should be carefully considered.
- Chronic toxicity studies are probably not required if a suitable microcosm or mesocosm study is available. However, it should be noted though that microcosm or mesocosm tests do not usually include the end-point of chronic toxicity to fish. Where valid fish data from a microcosm study (eg survival, growth, and behaviour) or mesocosm study (eg free living, reproduction data) are available then these may fulfill the requirement.

Simulated or actual field-testing

These are Tier III tests and may be conducted to quantify the actual risk in the field or to show that the risk in the field is different from what is observed in the laboratory. They include microcosm and mesocosm studies. The need for such tests is considered on an individual case basis.

6.4 ASSESSMENT OF SECONDARY EXPOSURE EFFECTS

6.4.1 INTRODUCTION

Experience has demonstrated that the need for an assessment of secondary exposure is uncommon. However, if there is the potential for a substance to bioaccumulate, a discussion on the possibility of adverse effects due to secondary exposure is recommended (OECD, 2004 (b)). Despite making this recommendation, there is currently no guidance at OECD level on how to assess the likely effects due to secondary exposure. The following discussion on assessment of secondary exposure follows methodology described in the corresponding industrial chemicals guidance assessment manual which is based on the TGD to maintain consistency.

The TGD (EC, 2003 (a)) provides guidance on undertaking an assessment of secondary exposure which is dependent on the bioaccumulation potential of the chemical. It employs equilibrium partitioning methodology in the initial screen. Assessors should be familiar with the guidance provided in this document, and the discussion in this section is paraphrased from the TGD unless otherwise indicated.

Assessors should also be familiar with the general concepts of bioaccumulation, and Appendix V provides a discussion on this end-point as described in the GHS. Essentially, if at the base set level, a substance:

- has a log Kow ≥3, or
- is highly adsorptive, or
- belongs to a class of substances known to have a potential to accumulate in living organisms, or
- there are indications from structural features, and
- there is no mitigating property such as hydrolysis (half life <12 h),

then there is an indication of bioaccumulation potential (EC, 2003 (a)).

In summary, the TGD explains the general approach to effects assessment for secondary exposure as:

- The assessment of the potential impact of chemicals on top predators is based on the accumulation of hydrophobic chemicals through the food chain. These may follow many different pathways along different trophic levels.
- In the absence of data on other uptake routes, it is assumed that direct uptake from water accounts for 100% of the intake and secondary exposure is not considered.
- For substances with a log Kow <4.5, the primary uptake route is direct uptake from the water phase. Where the log Kow >4.5, other uptake routes such as intake of contaminated food or sediment become increasingly important.

6.4.2 SECONDARY EXPOSURE THROUGH THE AQUATIC FOOD CHAIN

Assessment of risk to fish as a result of the combined intake of contaminants from water and contaminated food (aquatic organisms) is not considered necessary, as this is generally covered by the aquatic risk assessment and the risk assessment for secondary exposure of fish-eating predators. This section includes discussion of how to determine the PEC and PNEC relevant for assessing secondary exposure.

The risk to fish eating predators (mammals and/or birds) is calculated as the ratio between the concentration in their food (PECoral, predator) and the no effect concentration for oral intake (PNECoral) (this is the same as that used to evaluate direct effects on birds or mammals discussed earlier in this chapter). The concentration in fish is a result of uptake from the aqueous phase and intake of contaminated food. Therefore, the PECoral, predator is calculated from the bioconcentration factor (BCF) and a biomagnification factor (BMF). The PECoral, predator could also be calculated for other relevant species that are part of the food of predators.

$$PEC_{oral,\ predator} = PEC_{water} \times BCF_{fish} \times BMF$$

The BMF is the relative concentration in a predatory animal compared to the concentration in its prey. Where possible, the concentrations used to derive and report BMF values should be lipid normalised.

Foraging area will have an impact on the PEC estimation. If the local PECwater is used, this may overestimate the risk as it assumes predators will obtain 100% of their prey from the local area. However, the regional PECwater may lead to the opposite outcome as there may be areas with higher concentrations within the regional area. The TGD recommends a scenario where 50% of the diet comes from a regional area and 50% from the local area and this will define the PECwater used in the above calculation.

While the BMF should ideally be based on measured data, it is recognised such data are scarce. Therefore, default values are proposed in the TGD which assume a relationship between the BMF, the BCF and the Log Kow as follows:

Table 6-1: Default BMF values for organic substances

Log Kow of Chemical	BCF (fish)	BMF
<4.5	<2000	1
4.5 - <5	2000 - 5000	2
5 - 8	>5000	10
>8 - 9	2000 - 5000	3
>9	<2000	1

The PNECoral is ultimately derived from the toxicity data (dietary) applying an assessment factor (AF) as follows:

$$PNEC_{oral} = Tox_{oral}/AF_{oral}$$

The Toxoral is either a dietary LC50 bird, NOECbird or NOECmammal. Acute lethal doses for mammals and birds are not acceptable for extrapolation to chronic toxicity because these are not dietary tests. Conversely, acute effect concentrations for birds are acceptable for extrapolation. The results of available mammalian or avian tests may be expressed as a concentration in the food (mg/kg) or a dose (mg/kg body weight/day) causing no effect. Many standard tests aim to determine a dose, or a no-observed-adverse-effect-level (NOAEL), even when presented in food. Nevertheless, for the assessment of secondary poisoning, the results always have to be expressed as the concentration in food; therefore, dose concentrations should be converted to food concentrations. This information may be available from the studies provided body weights and daily food intakes are provided (to convert, the dose rate should be multiplied by the body weight/daily food intake).

Where this information is not available through studies, the TGD lists several mammalian and one bird conversion factor that may be used to calculate the NOEC from the dose level (NOAEL). In addition, the US EPA converts the residue concentrations to a daily oral dose, based on the fractions of body weight consumed daily (as estimated through mammalian allometric relationships). The US EPA's *Wildlife Exposure Factors Handbook* may provide surrogate information for this exercise (US EPA, 1993). On occasion, Australian wildlife data on body weights and daily food ingestion rates may be available in the literature.

The AF should compensate for specific aspects in the effects assessment of predators. A factor of 30 is considered to be appropriate for this purpose, accounting for both interspecies variation and lab-to-field extrapolation. Additionally, acute/sub-chronic to chronic extrapolation must be taken into account. The resulting AFs are given in the TGD and are reproduced below:

Table 6-2: Assessment factors for extrapolation of mammalian and bird toxicity data

Tox_{oral}	Duration of test	AF_{oral}
$LC_{50,\ bird}$	5 days	3,000
$NOEC_{bird}$	Chronic	30
$NOEC_{mammal}$	28 days	300
	90 days	90
	chronic	30

If a NOEC for both birds and mammals is given, the lower of the resulting PNECs is used in the risk assessment.

It should be recognised that this is a very simplistic process. Any information that may improve the input data or the assessment should be considered. If this assessment leads to the conclusion that there is a risk from secondary exposure, the option of undertaking additional laboratory tests (bioaccumulation in fish or feeding studies with laboratory mammals or birds) could be considered in order to obtain better data.

6.4.3 SECONDARY EXPOSURE THROUGH THE TERRESTRIAL FOOD CHAIN

Biomagnification may also occur via the terrestrial food chain. A similar approach as with the aquatic route is used, except fish are substituted with earthworms. The PNECoral is derived in the same way as the aquatic route (see above). Since birds and mammals consume worms including their gut contents, and the gut of earthworms can contain substantial amounts of soil, the exposure of the predators may be affected by the amount of substance in the soil. The PECoral, predator is calculated as:

$$PEC_{oral,\ predator} = C_{earthworm}$$

where:

Cearthworm is the total concentration of the substance in the worm as a result of both bioaccumulation into the worm tissues and the contaminated soil present in the gut.

As with the aquatic secondary exposure assessment, the concentrations in soil (Csoil) and porewater (Cporewater) are made on the assumption that 50% of the diet comes from the PEClocal and 50% from the PECregional.

The concentration in the earthworm (Cearthworm) is calculated as follows from the TGD:

$$Cearthworm = ((BCFearthworm \times Cporewater \times Wearthworm) + (Csoil \times Wgut))/(Wearthworm + Wgut)$$

where:

BCFearthworm = BCF for earthworms on wet weight basis (L/kgwet earthworm)

Wearthworm = weight of earthworm tissue (kg/kgwet tissue)

Wgut = Weight of gut contents (kg/kgww).

In turn, these parameters are calculated as:

$$W_{gut} = W_{earthworm} \times F_{gut} \times Conv_{soil}$$

where:

Convsoil = conversion factor for soil concentration wet-dry weight soil (kgww/kgdw)

Fgut = fraction of gut loading in worm (taken to be 0.1)

The conversion factor is necessary as the gut loading is determined in dry weight and is calculated as follows:

$$Conv_{soil} = BD_{soil}/F_{solid} \times BD_{solid}$$ (ie 1.25 using the MacKay default values)

where:

BDsoil = Bulk density of soil (1,500 kg/m3 from Mackay Level III Fugacity Model)

Fsolid = volume fraction of solids in soil (0.5 from Mackay Level III Fugacity Model)

BDsolid = Bulk density of solid (2,400 kg/m3from Mackay Level III Fugacity Model).

Where no bioconcentration data are available (as would usually be the case), bioconcentration can be calculated by the following QSAR described in the TGD:

$$BCF_{earthworm} = 0.84 + 0.012K_{ow}/BD_{earthworm}$$

where:

BDearthworm is the relative density of an earthworm and is assumed to be 1 kg/L.

This approach should be performed bearing in mind the following points from the TGD:

- This approach performed well in describing uptake in experiments with earthworms kept in water. For soil exposure, the experimental BCFs are generally somewhat lower than the model predictions, although the reasons for this discrepancy are unclear.
- Earthworms are also able to take up chemicals from food. While it has been hypothesised this process may affect accumulation at log Kow>5, data collected do not indicate this exposure route actually leads to higher body residues than expected on the basis of simple partitioning. Care must be taken in situations where the food of earthworms is specifically contaminated although reliable models to estimate this route are currently lacking.

The model was supported by data with neutral organic chemicals in soil with log Kow in the range 3-8 and in water only experiments with chemicals having a log Kow from 1-6. An applicable Kow range of 1-8 is advised and it is reasonable to assume that extrapolation to lower Kow values is possible. The underlying data are too limited at this stage to propose this approach in general for ionised chemicals.

CHAPTER 7 – ASSESSMENT OF PERSISTENT, BIOACCUMULATIVE AND TOXIC (PBT) SUBSTANCES

7.1 INTRODUCTION

Each aspect of persistence, bioaccumulation and toxicity has been considered separately throughout the normal risk assessment process (refer Chapters 5 and 6). However, it is often more difficult to estimate risks for a chemical that is classed as all three -- persistent, bioaccumulative and toxic. For this reason, such substances merit further consideration and the approach is outlined below.

7.1.1 INTERNATIONAL PERSPECTIVES

A survey on persistent, bioaccumulative, and toxic pesticides in OECD member countries was conducted during 1999-2000. The primary objective of the survey was to develop a clear understanding of: a) the information generally available to pesticide regulators that is relevant to risks associated with low-dose exposure to persistent, bioaccumulative, and toxic (PBT) pesticides; and b) how this information is used. The survey was undertaken with a view to developing a harmonised OECD approach for assessing the risks associated with exposure to these PBT pesticides that might be present at low levels in the environment. The outcomes of this exercise are reported in OECD, 2003 (b) and are summarised as follows:

Persistence: There is a lack of a harmonised definition for the persistence of pest control products among member countries. This appears to stem from a fundamental difference in defining persistence as either dissipation from a particular medium such as soil, or resistance to degradation. As a result, various numerical criteria are used to define persistence. For most countries, evaluation of the persistence of a pest control product requires flexibility in data requirements as well as the use of expert judgement in the interpretation of results. The assessment of the risk from major and minor metabolites differs among responding countries, where such an assessment is carried out. As well, the classification schemes used to assess persistence differ among countries. Persistence modeling is rarely used and monitoring for the persistence of pesticides is occasionally conducted on surface water, but less frequently on ground water.

Bioaccumulation: Almost all OECD countries examine bioaccumulation, however, the data are interpreted differently among countries. Bioconcentration factors, bioaccumulation factors, and octanol-water partition coefficients are used as indicators of bioaccumulation. Modeling for bioaccumulation of pesticides is not performed by any of the responding countries. Furthermore, monitoring for bioaccumulation is not typically conducted. Where biomagnification of pest control products is examined, various methods are used to assess the risk of biomagnification. Based on the received responses, there appears to be a need for harmonising both the definition of bioaccumulation and the approaches used in the assessment of bioaccumulation.

Toxicity: In general, the data requirements for the determination of toxicity are similar among the majority of responding countries. Most countries require or conditionally require studies on earthworms, bees, *Daphnia*, freshwater algae, fish, and birds. Different classification schemes may be used to assess toxicity in the same class of organisms among responding countries. Differences in the classification schemes include variations in the numerical end-points of the ranges for a particular species, the number of ranges used, and/or the descriptions of the degree of toxicity. The same classification scheme is, however, used for *Daphnia* and fish within each responding country. In most countries, modeling is not used to determine toxicity. As well, monitoring for toxicity is not required.

The majority of responding countries use models or calculations that account for multiple applications when estimating exposure. Finally, various approaches are used to examine the potential for biomagnification.

7.1.2 DATA GAPS AND OTHER APPROACHES

Although there are differences among responding countries in the types of data required and the approaches used to assess the submitted data, there were similar trends, viewpoints, and concerns identified by the responding countries.

Overall, there was consensus that studies on bioconcentration, bioaccumulation, and biomagnification should be required for the assessment of pest control products, however, the approaches to such assessments and their interpretation differ among responding countries. There was general agreement that certain circumstances exist under which persistent compounds should not cause environmental concern. These circumstances include reduced exposure, low toxicity, or a proposed use pattern that minimizes release to the environment. Numerous countries indicated that the persistence and toxicity of minor metabolites should be addressed in a separate manner from that of major metabolites. Relevance, toxicity, quantity, and risk were considered more important

than the designation of major or minor. Also, countries stated that there was a need for improving or furthering the science used in the assessment of persistence and toxicity.

Several important issues were identified by member countries as areas requiring improvement, especially in light of the ongoing harmonisation efforts. For the determination or assessment of persistence, the following issues were identified as areas needing further study or refinement:

1. agreement on a definition of persistence
2. harmonisation of approaches to the classification of persistence
3. establishment of clear criteria for defining persistence
4. use of modeling and measurement
5. determination of dissipation values
6. distinguishing between chemical persistence and biological persistence
7. improvement of test methods
8. assessment of persistence under the specific climatic conditions of each country
9. defining the role of bioavailability
10. development of methodology (exposure/risk assessment or toxicity tests) to assess the risk from prolonged exposure
11. defining the contribution of metabolites to exposure.

There is also a need to harmonise the definitions of bioconcentration/bioaccumulation and the approaches that are used in the assessment of these processes in risk assessments. Approaches to examining the potential for biomagnification also differ between member countries.

For the determination or assessment of toxicity, the following issues were identified as areas requiring improvement:

1. agreement on and determination of low level effects concentrations (ie ECx)
2. unification and updating of (test) methods
3. endocrine effects
4. chronic and/or population effects
5. teratogenicity
6. long-term risk assessment for mammals and guidance regarding the selection of the appropriate, ecologically-relevant end-points from mammalian toxicity tests
7. harmonisation of decision making on hazard assessment
8. examination of the link between laboratory tests and field tests.

Harmonisation of data requirements and scientific approaches used in the assessment of pest control products will facilitate the exchange of pesticide evaluations among OECD member countries. Such efforts will reduce the amount of work required in the scientific review process, thus facilitating the registration and re-registration of pest control products. Furthermore, harmonisation of data requirements and approaches will reduce trade barriers between countries, as well as reduce the need for duplicative testing of pesticides by industry.

Assessors should be aware of developments in this area. As an interim approach to PBT assessment, the procedures recommended for PBT assessment of industrial chemicals (Lee-Steere, 2007) are proposed in the following sections for use with agvet chemicals.

7.1.3 Conclusion

There are many similarities in the data requirements and the method of scientific review of the data for persistent, bioaccumulative, and toxic pesticides among responding OECD member countries. However, there are several areas of discrepancy in the interpretation of these data requirements. The ways in which these data are interpreted and used differ as a result of the regulatory approaches and policies of each country. These differences may lead to unique regulatory actions in each country for the same pest control product. Nevertheless, future efforts to increase the international harmonisation of data requirements should be enhanced

by the inventory of current national data requirements for persistent, bioaccumulative, and toxic pesticides developed from the survey described above.

7.2 OVERVIEW

As we know, it may be more difficult to estimate risks posed by a chemical that is classed as persistent, bioaccumulative and toxic, particularly where data packages are relatively small and more estimation is required. This is because traditional risk assessment methodologies may not adequately address the following concerns:

a) The concern that hazardous substances may accumulate in parts of the environment and that:
   i. The effects of such accumulation are unpredictable in the long-term
   ii. That such accumulation would be practically difficult to reverse.

b) The concern that remote areas of the oceans should remain untouched by hazardous substances resulting from human activity, and that the intrinsic value of pristine environments should be protected.

As well, these chemicals can often travel long distances. This means that it is not often possible to control their movement into another country and it is also not possible to properly inform neighbouring countries that might be affected by the movement of these chemicals.

These concerns occur with substances that persist for long periods, bioaccumulate in biota and can give rise to toxic effects after a greater time and at a greater distance than chemicals without these properties. Because of these properties, the TGD predominantly focuses on the marine environment since once the chemical has entered the open seas, any cessation of emission will not necessarily result in a reduction in chemical concentration. Consequently, any effects become particularly difficult to reverse. In addition, because exposures can be long-term and because many important marine species have a long life cycle, effects may be difficult to detect at an early stage. Correspondingly, a 'safe' concentration is difficult, if not impossible, to establish.

In undertaking a PBT assessment, it is first necessary to identity PBT substances using specific criteria for the inherent properties of the chemical. DEWHA has examined the existing criteria for both POPs and PBT substances, and proposed criteria have been adopted in Australian environmental risk assessments of chemicals, and these criteria are provided below. In line with other countries or regional approaches two tiers are proposed, one for persistent and bioaccumulative substances, and the other for very persistent and very bioaccumulative substances. These criteria can be used both for preventing new pesticides/veterinary medicines or industrial chemicals exhibiting POPs characteristics from being placed on the market, or for screening of these chemicals for priority setting of future assessments. It should be noted that the Stockholm Convention criteria are in the context of POPs, therefore, can be defined as being for *very* persistent and *very* bioaccumulative substances. A chemical deemed persistent or bioaccumulative may not carry values as high as those prescribed in the POPs criteria.

Chemicals that are found to be persistent, bioaccumulative and toxic are not always amenable to an RQ approach and must be handled on a case-by-case basis in consultation with the regulatory agencies.

7.3 DATA EVALUATION AND AVAILABLE GUIDANCE

Procedures for determining the quality of data, use of analogue data, use of QSARs and application of expert judgement should still be undertaken in accordance with Chapter 3.

In addition to the information provided in Chapter 3 as well as this chapter, guidance documents are available internationally for assisting in assessing PBT characteristics of a chemical.

The Canadian guidance manual for the categorisation of substances on their domestic substances list (Environment Canada, 2003) provides guidance for determining the persistence, bioaccumulation potential and inherent toxicity to non-human organisms.

The US EPA has developed the PBT profiler <www.pbtprofiler.net> to predict the PBT potential of chemicals based on USA criteria. Also, the TGD provides additional guidance on undertaking a PBT assessment.

7.4 PBT CRITERIA

7.4.1 PERSISTENCE CRITERIA

The Stockholm Convention provides scientifically based criteria for potential POPs and a process that ultimately may lead to elimination of a POP substance globally. The criteria for persistence in Annex D of the convention are expressed as **single-media** criteria as follows:

i. Evidence that the half-life of the chemical in water is greater than two months, or that its half-life in soil is greater than six months, or that its half-life in sediment is greater than six months, or

ii. Evidence that the chemical is otherwise sufficiently persistent to justify its consideration within the scope of the Convention.

Considerations under the second criteria may include, for example, such situations as where the chemical has a half-life greater than six months in anaerobic conditions, but less than six months under aerobic conditions.

The following persistence criteria have been adopted by DEWHA with definitions from the Stockholm Convention remaining for very persistent compounds.

Persistent (P)	
For PBT purposes a chemical is considered persistent in a particular media if its half life in the media exceeds the following:	
Media	**Half-Life**
Water	2 months
Soil	6 months
Sediment	6 months
Air	2 days

Criteria prescribed within the EU, United States and Canada are summarised as follows:

European Union (EC, 2003 (a))

Persistent and very persistent substances (EU)

Persistent (P)	Very persistent (vP)
Half-life >60 d in marine water or >40 d in freshwater or half-life >180 d in marine sediment or >120 d in freshwater sediment[1)]	Half-life >60 d in marine- or freshwater or >180 d in marine or freshwater sediment

Overview of P-assignment for different types of biodegradation data (EU)

Type of data	Criterion	Definitive assignment	Screening assignment[1)]
DT50 marine water	> 60 d	vP	-
DT50 freshwater[2)]	> 40 d	P [3)]	-
	> 60 d	vP	-
DT50 marine sediment	> 180 d	vP	-
DT50 freshwater sediment	> 120 d	P [3)]	-
	> 180 d	vP	-
Readily biodegradable [4)]	Yes	Not P	-
	No	-	P or vP
Inherently degradable	Yes	Not P [5)]	-
	No	-	P or vP
QSAR	Thresholds defined for different models.	-	P or vP

1) These screening methods give an "open-ended" categorisation of the substance as either being potentially P or vP, which cannot be related to a half-life for biodegradation.

2) Data for estuaries should also be considered in this category.

3) Half-life data in freshwater and freshwater sediment can be overruled by data obtained under marine conditions.

4) Regardless of whether the 10-d window criterion is fulfilled.

5) This only applies to cases where the specific criteria as mentioned in Section 4.4.3.3. are fulfilled.

United States of America (US EPA, 1999)

	Considered persistent	Considered very persistent
Half-life in water, soil, and sediment	Half-life >= 2 months (>= 60 days)	Half-life > 6 months (> 180 days)
Half-life in Air	Half-life > 2 days	

Canada

Persistence criteria for Canada were prescribed in the Canada Gazette, 2000 as follows:

Persistence [1]	
Medium	Half-life
Air	≥2 days [2]
Water	≥6 months
Sediment	≥1 year
Soil	≥6 months

1) A substance is considered persistent when the criterion is met in any one medium.

2)A substance may be considered as persistent in air if it is shown to be subject to atmospheric transport to remote regions such as the Arctic.

7.4.2 BIOACCUMULATION CRITERIA

As noted above, the Stockholm Convention on POPs provides scientifically based criteria for potential POPs and a process that ultimately may lead to elimination of a POP substance globally. The criteria for bioaccumulation in Annex D of the convention are given as follows:

i. Evidence that the bioconcentration factor or bioaccumulation factor in aquatic species for the chemical is greater than 5000 or, in the absence of such data, that the log Kow is greater than 5

ii. Evidence that a chemical presents other reasons for concern, such as high bioaccumulation in other species, high toxicity or ecotoxicity, or

iii. Monitoring data in biota indicating that the bioaccumulation potential of the chemical is sufficient to justify its consideration within the scope of the Convention.

It should be noted that POPs are defined as being very persistent and very bioaccumulative substances. A chemical deemed persistent or bioaccumulative may not carry values as high as those prescribed in the POPs criteria.

The following bioaccumulation criteria have been adopted by DEWHA with definitions from the Stockholm Convention remaining for very bioaccumulative compounds.

Bioaccumulative (B)
For PBT purposes a chemical may be considered to be Bioaccumulative if it has a BCF/BAF or >2000, or in the absence of any BCF/BAF measurements, a logKow >4.2

As before, criteria prescribed within the EU, United States and Canada are summarised as follows:

European Union (EC, 2003a)

Bioaccumulative and very bioaccumulative substances (EU)

Bioaccumulative (B)	Very bioaccumulative (vB)
BCF >2,000	BCF >5,000

United States of America (USEPA, 1999 <http://www.epa.gov/fedrgstr/EPA-TOX/1999/November/Day-04/t28888.htm>

Bioaccumulative (B)	Very bioaccumulative (vB)
BCF >1,000	BCF >5,000

Canada (Canada Gazette, 2000)

Bioaccumulation
BAF ≥5,000 Or BCF ≥5,000 Or Log K_{ow} ≥5

For a more detailed discussion on assessing the bioaccumulation potential of a chemical, refer to Appendix V.

7.4.3 TOXICITY CRITERIA

Although toxicity has been considered in the previous chapter, for persistent and bioaccumulative substances, exposure may be anticipated to cover the whole life of an organism as well as multiple generations. Consequently, chronic ecotoxicity data, preferably covering impacts on reproduction, should ideally be used to establish the toxicity within the PBT context. DEWHA uses a range of international values in a standard approach, as detailed below.

As noted above, the Stockholm Convention on POPs provides scientifically based criteria for potential POPs and a process that ultimately may lead to elimination of a POP substance globally. The criteria for toxicity in Annex D of the convention do not consist of numerical values, but are given as follows:

(e) Adverse effects:

(i) Evidence of adverse effects to human health or to the environment that justifies consideration of the chemical within the scope of this Convention; or

(ii) Toxicity or ecotoxicity data that indicate the potential for damage to human health or to the environment.

The following toxicity criteria have been adopted by DEWHA.

Toxic (T)		
For PBT purposes, in respect of aquatic toxicity, a chemical may be considered toxic under the following circumstances (corresponding to criteria for GHS chronic category 1):		
Non-rapidly degradable substances for which there are adequate chronic toxicity data available	Chronic NOEC or EC_x (for fish)	≤0.1 mg/L and/or
	Chronic NOEC or EC_x (for crustacea)	≤0.1 mg/L and/or
	Chronic NOEC or EC_x (for algae or other aquatic plants)	≤0.1 mg/L
Rapidly degradable substances for which there are adequate chronic toxicity data available	Chronic NOEC or EC_x (for fish)	≤0.01 mg/L and/or
	Chronic NOEC or EC_x (for crustacea)	≤0.01 mg/L and/or
	Chronic NOEC or EC_x (for algae or other aquatic plants)	≤0.01 mg/L
Substances for which adequate chronic toxicity data are not available (providing criteria for P and B are met)	96 h LC_{50} (for fish)	≤1 mg/L and/or
	48 h EC_{50} (for crustacea)	≤1 mg/L and/or
	72 or 96 h ErC_{50} (for algae or other aquatic plants)	≤1 mg/L
	And the substance is not rapidly degradable and/or the experimentally determined BCF is ≥500 (or, if absent, the logKow is ≥4	
Toxicity to other (terrestrial) organisms	Should be considered on a case by case basis, compared with the highly toxic classifications DEWHA has developed for agvet chemicals	
Long term toxicity or evidence such as endocrine disruption effects	Should be considered on a case by case basis.	

Similarly, criteria prescribed within the EU, United States and Canada is summarised as follows:

European Union (EC, 2003a)

Toxicity (T)
Chronic NOEC <0.01 mg/L or CMR[1] or endocrine disrupting effects.

1) CMR = Carcinogenic, mutagenic or toxic to reproduction

United States of America

Unlike persistence and bioaccumulation, US EPA, 1999 does not provide numerical criteria for toxicity values in legislation. As explained in this reference, a number of submissions contended that the EPA should set a separate toxicity criteria for PBT chemicals. The EPA disagreed. Their EPCRA (Emergency Planning and Community Right to Know Act) Section 313 provides toxicity criteria at Section 313(d)(2) to be used in adding a chemical to or deleting a chemical from the EPCRA Section 313 list of toxic chemicals. These criteria with respect to the environment are:

> The chemical is known to cause or can reasonably be anticipated to cause, because of-
>
> (i) its toxicity
>
> (ii) its toxicity and persistence in the environment, or
>
> (iii) its toxicity and tendency to bioaccumulate in the environment
>
> (iv) a significant adverse effect on the environment of sufficient seriousness, in the judgement of the Administrator, to warrant reporting under this section.

Rather, to highlight a chemical that may be chronically toxic to fish, the PBT profiler uses criteria developed in EPA's new chemical program <http://www.pbtprofiler.net/criteria.asp>. These criteria are:

	Low concern	Moderate concern	High concern
Fish ChV[1] (mg/L)	>10 mg/L	0.1-10 mg/L	<0.1 mg/L

1) ChV = chronic value, or MATC

Canada (Environment Canada, 2003)

Environment Canada provides the following discussion and classification criteria for inherent toxicity to non-human organisms.

In toxicology terms, *toxicity* is the inherent potential or capacity of a material to cause adverse effects on living organisms. Consequently, in common use, *toxic* means 'able to cause injury to living organisms as a result of physicochemical interaction'

Environment Canada uses the term *inherent toxicity* to distinguish from the word toxic. As explained, *inherent toxicity* refers to the hazard a substance presents to the environment or human health, which can be represented by the toxic effect caused by the substance - ie the toxicity found in a study or predicted due solely to the test substance, or the effect that has not been masked or mitigated by some factor or parameter.

The categorisation of substances for inherent toxicity should ideally use both aquatic (including benthic) and terrestrial species. However, the overwhelming majority of experimental ecotoxicological data are obtained in tests with aquatic/pelagic species. In addition, most of the LC50 and Kow values available for categorisation are based on model prediction. Virtually all of the quantitative structure–activity relationship (QSAR) estimates (as well as experimental toxicity data) have been generated employing external effect concentrations in the aquatic environment. Consequently, the aquatic compartment will be used systematically to categorise the substances.

Environment Canada has reviewed the current science concerning inhalation toxicity tests and data. No recognised standard tests/methods on inhalation toxicity for invertebrates, amphibians, reptiles, or birds are available at present. Furthermore, as almost all existing inhalation toxicity data refer to mammalian toxicity, Environment Canada will consult with Health Canada, which will review mammalian toxicity data.

The categorisation for inherent toxicity is based on numerical criteria (see below). When reliable results on chronic studies are available, the chronic toxicity values will be applied.

Table 7-1: Criteria for acute and chronic toxicity to aquatic species (algae, invertebrates, fish)

Exposure duration	Criteria
Acute	LC_{50} (EC_{50}) =1 mg/L
Chronic	NOEC <0.1 mg/L

The above approach proposed by Environment Canada is in agreement with some international initiatives, such as the OECD's screening information data set (SIDS). Acute toxicity data for fish, *daphnia*, and algae are required elements in the ecotoxicity section of the SIDS data elements. Other data, such as terrestrial toxicity data, are not systematically sought.

7.4.3.1 Determining the toxicity of a chemical

Mammalian toxicity data should also be considered in addition to aquatic toxicity data, because toxic effects on top predators, including man, may occur through long-term exposure via the food chain. The selection criteria should therefore consider two types of effect data (chronic or acute), either of which will trigger selection.

Chronic effects data

A substance is considered to fulfil the toxicity criterion when:

- the long-term NOEC for marine or freshwater organisms is less than the trigger value. When other information is available such as data on sediment toxicity or data from feeding studies, this needs to be assessed on a case-by-case basis. Results from subchronic, chronic or reproduction avian toxicity tests may be available for biocides and pesticides. The TGD suggests a chronic NOEC of <30 mg/kg food be used as a trigger criterion.

or

- when the substance is classified as carcinogenic, mutagenic or toxic for reproduction, or when there is evidence of chronic toxicity. In these cases, assessment must be carried out to decide whether the evidence is sufficient for the substance to be considered as toxic, in the context of the PBT assessment, or whether further information is needed to clarify this potential concern.

or

- when there is substantiated evidence of long-term toxicity (eg.endocrine disrupting effects). Such evidence needs to be considered on a case-by-case basis.

Acute effects data (screening level)

Where data on chronic effects are not available then short-term toxicity data for marine or freshwater organisms can be used to determine whether a substance is a potential PBT, provided the screening criteria for P and B are fulfilled. In PBT assessment, trigger values for the T criterion are not provided for acute data. The TGD suggest a substance is considered potentially toxic when the L(E)C50 to aquatic organisms is <0.1 mg/L. If a substance fulfils the ultimate P and B criteria, chronic toxicity data would be required to deselect this substance from being considered PBT. In principle, chronic toxicity data should override the results from the acute tests, when obtained for the same species.

Acute mammalian toxicity tests are not normally considered to provide an appropriate indication of chronic effects in the context of the PBT-assessment. However, it should be noted that when a substance is classified as very toxic or toxic after oral dosing (LD50 <200 mg/kg bw/d) and toxicity is expected to be the result of systemic effects, the probability that the chronic NOAEL after repeated dosing will be less than the (EU) trigger value will be high. The substance would therefore be classified and considered as fulfilling the T-criterion. In that case verification of the actual chronic toxicity by performing animal testing is not recommended. When the P and B screening criteria are also fulfilled, the substance can be considered as a PBT unless additional information indicates otherwise.

Estimated effects data

Where no acute or chronic toxicity data are available, the assessment of the T-criterion at a screening level can be performed using data obtained from QSARs.

As noted in the TGD, further toxicity testing for very bioaccumulative substances, is necessary since long term effects can be anticipated.

In conclusion, substances fulfilling PBT criteria are of priority for further consideration with the ultimate goal to restrict, if not end, any emissions to the environment. For such substance, an evaluation of the sources, major emissions and pathways to the environment should take place in order to sufficiently establish the most appropriate and effective measures to reduce releases to the environment.

CHAPTER 8 – RISK CHARACTERISATION

8.1 INTRODUCTION

Risk characterisation is the final phase of the ecological risk assessment where the outcomes of the exposure (Chapter 5) and effects assessments (Chapter 6) are integrated. In the deterministic risk assessment approach used by Australia, the primary outcome of the risk characterisation is the calculation of the risk quotient (RQ). The RQ is compared with levels of concern in order to draw a conclusion regarding risk. This conclusion may range from acceptable risk through to the requirement for further testing or risk management options to reduce risk to acceptable levels, or to unacceptable risk.

When reporting the RQ, a description of the uncertainties, assumptions, strengths and limitations associated with it should also be provided. These will largely be discussed during characterisation of the exposure and effects and will include refinement options used in ultimately determining the RQ. Such refinement options are discussed in this chapter, along with the end-points and values to consider when calculating an RQ for the following taxanomic groupings:

- Birds and mammals (Section 8.1)
- Invertebrates such as beneficial insects (Sections 8.2-8.4)
- Soil micro-organisms (Section 8.5)
- Terrestrial plants (Section 8.6), and
- Aquatic organisms (Section 8.7)

As well, consideration is given to levels of concern (Section 8.8), and risk management options (Section 8.9).

8.2 BIRDS AND MAMMALS

Calculation of the RQ should be determined using the lowest reliable value for the end-point being considered, that is, acute LD50, dietary LD50 or chronic NOEC. Where data are available for both the active constituent and the formulation, the exposure context should be considered. Once a determination has been made regarding the more appropriate exposure route (eg formulation in the case of granules or active substance with spray formulations) then the relevant figure should be used.

Some formulae below refer to a peak nomogram concentration. Refer to Section 5.4.1 for the discussion on obtaining nomogram concentrations.

RQs are determined for acute scenarios, dietary scenarios and chronic scenarios, as follows:

8.2.1 ACUTE (SCREENING LEVEL)

Granular formulation mg ac/m2/Most sensitive bird oral LD50 = RQ

The outcome of this formula will enable a qualitative comment regarding the area a bird will need to forage in order to receive a lethal dose and should be considered in interpreting the RQ.

8.2.2 DIETARY (SCREENING LEVEL)

Peak nomogram concentration for diet/Most sensitive bird dietary LC50 = RQ

8.2.3 CHRONIC (SCREENING LEVEL)

Peak nomogram concentration for diet*/most sensitive bird reproduction NOEC = RQ

*Multiple application peaks are added.

8.2.4 CHRONIC (REFINED)

Peak nomogram concentration for diet**/most sensitive bird reproduction NOEC = RQ

**Assessment considers pesticide dissipation for multiple application scenarios. Risk characterisation may include discussion of the duration of exposure above the toxicity thresholds.

8.3 BEES

8.3.1 STANDARD RISK ASSESSMENT

The maximum single application rate is used to determine the RQ to bees. A bee is assumed to be around 1 cm2 in size. The application rate is converted to the amount of active constituent per square centimetre (ac/cm2) and then compared directly to the oral and contact LD50 values as follows:

Application rate (g ac/cm2)/oral or contact LD50 = RQ

The outcome of these RQ values determines the need for further higher tier testing in order to adequately evaluate the risk to bees.

8.3.2 HIGHER TIER RISK ASSESSMENT

Expert judgement is needed to interpret the results from higher tier testing, because such testing is often specific to the substance and use scenarios in question. The following is paraphrased from EC, 2002 (b).

There are no clearly defined end-points for higher tier studies. Therefore, a degree of expert judgement is required to interpret both semi-field and field study results. In the former case - where there are replicated studies - there should be a statistical comparison between key parameters such as foraging density, mortality, and proportion of adults, larvae and pupae in the hive. Of course, the parameters considered should be relevant to the compound under assessment. For example, if an insect growth regulator were being assessed then it would be more relevant to concentrate on developmental measurements and end-points.

When assessing the results of field trials, key parameters should be compared to either pre-treatment levels or to control levels. It is important to consider any effects observed in relation to the overall survival and productivity of the hive.

Key parameters that may be considered in a field trial include: mortality (assessed via the use of dead bee traps), behaviour (including foraging behaviour in the crop and around the hive), honey crop (assessed via weighing the hive at appropriate intervals) and state of colony (including an assessment of brood). Depending upon the concern highlighted in the initial risk assessment it may be appropriate to use pollen traps as well as analysis of dead bees. Analysis of honey and wax may also be useful in determining exposure.

As with many ecotoxicity tests, the use of a toxic standard (or positive control) in both semi-field and field trials along with an untreated control will often aid interpretation of the results. For insect growth regulators and other active substances that may cause long-term adverse effects on hive health, data over a long time period is required in order to adequately assess effects. For a conclusion of acceptable risk to be drawn then evidence is required confirming a lack of effects on hive health over a long time period. Further information is available in the EPPO guideline (EPPO, 2001). The design of higher tier studies is dependent upon the risks that have been highlighted. Therefore it is recommended that applicants/registrants consult the relevant authority when designing such studies.

8.4 OTHER ARTHROPODS

8.4.1 CURRENT APPROACH

Exposure is directly related to the maximum application rate for both in-field (direct application) and off-field (through spray drift) exposure. In this case, an RQ is not specifically calculated; rather the results are interpreted based on IOBC classification (Boller *et al*, 2005) and the arthropod studies. For example, if compared to a control group, a mortality/reduction in beneficial capacity in parasitic mites in a laboratory study of 60% were found at the application rate proposed in Australia, the IOBC classification of 'moderately harmful' would apply.

Where an application rate is likely to result in a 'harmful' classification (ie, greater than 80% mortality/reduction in beneficial capacity), then qualitative arguments are considered. These may include out-of-crop refuge areas and in-field untreated areas if application is by band/row/furrow methods. Alternatively, higher tier studies may be requested.

8.4.2 EU APPROACH

The EU describes a more formal process for considering risk to non-target arthropods and this is described below (EC, 2002 b). This process relates to EU specific data requirement in this area, described above in Section 6.2.3.3:

8.4.2.1 Assessing the risk in-field

Step 1: Tier I assessment based on standard tests

In the first tier, the risk is characterised by the in-field hazard quotient (HQ):

In-field HQ = in-field exposure/LR50

where the LR50 comes from glass-plate tests with the two standard species. If the in-field HQ is less than 2 for both species, then no further assessment is required (for the reasoning behind this trigger level refer ESCORT 2, Candolfi *et al*, 2001). If the HQ is greater than or equal to 2 for one or both species, then proceed to Step 2.

Step 2: Higher tier assessment

If no appropriate risk mitigation measures can be identified, then the applicant/registrant should carry out higher tier studies on the affected species as well as one further species with different biology.

Details of suitable species are provided in ESCORT 2. Lethal and sub-lethal effects of less than 50% are considered acceptable in extended laboratory tests and semi-field tests, provided that the tests covered the appropriate field rate. Refer ESCORT 2 for interpretation of field studies and recolonisation in aged residue studies. In the latter case, it has to be demonstrated that there is a potential for recolonisation/recovery at least within one year but preferably in a shorter period depending on the biology (seasonal pattern) of the species. The assessment may be based on field studies or other evidence (eg results of aged-residue studies, environmental fate information). As always, any data and assumptions should be fully justified.

8.4.2.2 Assessing the risk off-field

Step 1: Tier I assessment based on standard tests

The risk is characterised by the 'off-field' HQ:

Off-field HQ = (off-field exposure/LR50) × correction factor

where the LR50 comes from glass-plate tests with the two standard species. The correction factor is intended to cover uncertainty regarding species sensitivity. The default value of the correction factor is 10. If the off-field HQ is less than 2 for both species, then no further assessment is required. As before, if greater than or equal to 2 for one or both species, then proceed to step 2.

Step 2: Higher tier assessment

If no appropriate risk mitigation measures can be identified, then higher-tier studies (see below) on the affected species as well as two additional species with different biologies should be conducted.

8.4.2.3 Higher tier testing

Once again, details regarding suitable species are provided in ESCORT 2. Lethal and sub lethal effects of less than 50% are considered acceptable for extended laboratory tests and semi-field tests, provided that the tests covered the appropriate field rate. In this case the default value for the correction factor is 5. Generally, it has to be demonstrated that there is an acceptable potential for recovery within an ecologically relevant period.

If the Tier 1 assessment indicates a risk then either risk mitigation measures or higher-tier studies are called for. Judgement must be applied as it is not considered appropriate to propose unrealistic risk mitigation measures (eg exaggerated buffer zones) in order to avoid higher-tier testing.

This standard approach is not appropriate for substances with limited solubility or for plant protection products such as granules, seed treatments and pellets. As well, it may not be wholly appropriate for insect growth regulators or other compounds with particular modes of action. In the former cases it is recommended that studies are conducted with *Hypoaspis aculeifer* or *Folsomia candida* as proposed by EPPO, 2002. If deemed appropriate, then studies with *Aleochara sp.* might be conducted, such as at Tier II. For insect growth regulators or compounds with particular modes of action, the principles of ESCORT 2 should be followed with case-by-case modification according to the specific issues for the compound in question.

8.5 EARTHWORMS AND OTHER SOIL MACRO ARTHROPODS

8.5.1 STANDARD RISK ASSESSMENT

The standard risk assessment is based on RQ values obtained through comparing the PECsoil with the LC50 from an acute earthworm study, or the accumulated PECsoil with the NOEC from a chronic earthworm study.

As explained in EC, 2002 (b), the toxicity of lipophilic organic contaminants to soil organisms usually depends on the organic carbon content (foc) of the substrate/soil as this governs adsorption and, thus, pore water

concentration. The artificial substrate of the earthworm laboratory tests has a higher foc than many natural soils, so it could be expected that the LC50 or NOEC would be lower if the test were conducted in natural soil.

The EU risk assessment method accounts for this difference by dividing the LC50 and the NOEC by 2 where the log Kow is greater than 2 (EPPO 2002) unless it can be demonstrated by soil sorption data or other evidence that the toxicity is independent of foc.

8.5.2 REFINED RISK ASSESSMENT FOR EARTHWORMS/SOIL MACRO INVERTEBRATES

If the acute or chronic RQs are above the levels of concern, then further action is required, either by refining the effects or exposure estimates, or requiring higher tier studies. EC 2002 (b) provides some guidance in these areas, as outlined in the following sub-sections.

8.5.2.1 Refined effects assessment

When the NOEC from a reproductive test is expressed in g/ha, it can be converted into mg/kg soil by a standard calculation. This standard calculation assumes 100% of substance reaching the soil, a 5 cm depth and a soil density of 1.5 in order to give a value used in the chronic RQ calculation. (**NOTE:** Australian practice is to use a 10 cm soil depth unless the chemical is immobile, then 5 cm is used. For more mobile chemicals, a soil depth of 15 cm can be considered) However, this calculation can be refined when RQ is close to the level of concern, by considering actual test values (application rate and surface of the test unit, dry soil weight in the test unit). If there are uncertainties arising from the fact that the standard tests are conducted with artificial soil, then an option might be to do the earthworm test in natural soil.

8.5.2.2 Refined exposure estimate

The exposure assessment could be improved by employing more sophisticated models, by consideration of interception, or by inclusion of field measurements.

8.5.2.3 Higher tier studies

Based on experience, it is becoming widely accepted that earthworms are much more sensitive to pesticides during longer term testing than they are in the acute 14 day test. Even where the acute RQ is below the level of concern, the earthworm reproduction test can be regarded as the next higher tier.

Current requirements for such a test are triggered by exposure considerations such as repeated applications or persistence properties of a chemical. The earthworm reproduction test fulfils two purposes. Firstly, it is a long-term test with sub-lethal end-points and secondly, it involves more realistic conditions, such as surface application instead of mixing into the soil, than does the acute test.

8.6 SOIL MICRO-ORGANISMS

An RQ is not determined when assessing risk to soil micro-organisms. Rather, the outcome of the soil micro-organism test is directly assessed in terms of risk. A deviation of 25% from the control in terms of respiration rates or nitrogen transformation is considered the critical level in ascertaining risk. Larger deviations will require a refinement of the assessment. The concentrations used in the test should at least be as high as the highest proposed application rate.

8.7 NON-TARGET TERRESTRIAL PLANTS

8.7.1 INTERNATIONAL BACKGROUND

Section 5.4.2 described an international workshop convened by US EPA on assessing the ecological risk of chemicals to non-target plants. The conclusions of this workshop with regard to risk characterisation of non-target terrestrial plants were as follows:

> Deterministic approaches present the risk as a quotient (RQ), based upon effects information for the most sensitive species tested. However, Europe is moving away from a single-species deterministic range of quotients. In addition to a quotient, other information that should be considered in risk characterisation includes: a range of toxicity data; slope of dose-response curve; nature of the effects end-point; frequency and timing of exposure event(s); recovery; scale and magnitude of product use; bioavailability; environmental fate of product; incident data; climate differences; and vulnerable habitats.
>
> Uncertainty can be expressed through the use of safety factors, but these are not used at all in some countries, including Australia. Uncertainty factors have been largely used to account for differences in species sensitivity, but could also be used to account for routes of exposure that were not considered, for regional variation, or for the uncertainty in fate data. It is thought that large uncertainty factors may be too conservative in some cases.

Probabilistic assessment may or may not reduce uncertainty, depending upon the accuracy of the input information.

In the absence of data on recovery, then assumptions of no recovery must be used. Recovery, if known, could be used in developing mitigation measures or for refining risk.

Monitoring data can be quite useful. Such data are used to reduce uncertainty; to refine assumptions; to verify the effectiveness of mitigation; to obtain a conditional registration; or to change a registration decision.

8.7.2 CURRENT APPROACH

In Australia, risk characterisation for non-target terrestrial plants is based on the most sensitive toxicity end-point (eg plant height, dry weight, etc.) using an LC50 if mortality is the end-point, or an EC25 based on adverse effects. Use of the latter, more conservative, value accounts for the longer regeneration times of many terrestrial plants. As noted in MORAG, the risk quotient (RQ) is calculated as follows:

RQ = PEC/toxicity value.

The end-points measured in most screening studies, such as phytotoxicity, chlorosis, etc., cannot be interpreted as a NOEC value covering germination and biomass production. However, the available information usually allows the use of a conservative approach, assuming, for example, that when an untreated control has been run in parallel, then any effect accounting for at least 25% reduction in biomass production could be identified in a visual inspection. In addition, single dose experiments reported in terms of percentage of observed effects can also provide indications on the potential hazard of the substance for terrestrial plants.

Higher tier risk assessment based on field studies

If likely environmental concentrations are above the EC25 then higher tier testing may be required. Analysis for a higher tier risk characterisation is done on a case-by-case basis. Key elements for the assessment include the ecological relevance of the observed effects, consequences on soil functions, and the potential for recovery (EC, 2002 (b)). It is recommended that applicants/registrants consult the relevant authority when considering conducting such studies to ensure the design is appropriate.

8.8 AQUATIC ORGANISMS

8.8.1 CURRENT APPROACH

The DEWHA uses the quotient approach (Urban and Cook, 1986) to predict environmental risk for aquatic organisms such as fish, aquatic invertebrates, aquatic plants and algae. The quotient (Q) of the predicted environmental concentration (PEC) and the relevant toxicity concentration (lethal concentration for 50% of a population, LC_{50} or lowest-observed-effect concentration, LOEC) acts as a trigger for further scrutiny. For LC_{50} or EC_{50} data, $Q>0.5$ results in a "presumption of unacceptable risk" to organisms, while if $0.1 \leq Q \leq 0.5$, there is a "presumption of risk that may be mitigated by restricted use". A value of $Q<0.1$ indicates a low potential environmental risk. A result of $Q<1$ is considered acceptable when a chronic NOEC is used as the effect end-point.

Calculation of an acute RQ for aquatic organisms:

PECwater/most sensitive organisms LC/EC50 = RQ

Calculation of a chronic RQ for aquatic organisms:

PECwater/ most sensitive organism NOEC = RQ

Note 1: The PECwater is the same value used for both acute and chronic RQ derivations. This differs from the USA where the 21 day average water concentration is used for chronic invertebrate risk characterisation while the 56 or 60 day average water concentration is used for chronic fish risk characterisation.

Note 2: The EU also uses bioconcentration factors (BCF) as triggers in the aquatic risk assessment. Authorisation in the EU will not be granted on BCF grounds where 'the maximum bio-concentration factor is greater than 1000 for plant protection products containing active substances which are readily biodegradable, or greater than 100 for those which are not readily biodegradable, unless it is clearly established through an appropriate risk assessment that under field conditions no unacceptable impact on the viability of exposed species (predators) occurs - directly or indirectly - after use of the plant protection product according to the proposed conditions of use'.

8.8.2 SEDIMENT ORGANISM CONSIDERATIONS

Where exposure may result from spray drift, or the soluble phase of a chemical in run-off, the risk characterization may be performed from results of a 'spiked water' study with sediment dwellers compared to surface water PEC values.

Where measured values are available, toxicity data from a 'spiked sediment' study should be compared with whole-sediment PEC values. This is an example of how exposure to sediment organicms may occur in the event of a soil bound chemical reaching the sediment compartment following run-off. Where measured values are not available for the PEC and/or PNEC in this case, a risk characterization of the sediment compartment is not feasible. If equilibrium partitioning has been applied to derive both the PEC and the PNEC, then the result will simply be as for the aquatic compartment.

8.8.3 HIGHER TIER RISK ASSESSMENT

The inclusion of data for more than the minimum number of species does not allow a reduction in the assessment factors described here to account for the lower uncertainty, or the increased weight of evidence associated with such increased data sets. In this respect the risk quotient approach described above does not allow much flexibility in refining effects where no higher tier studies are available. Therefore, where an RQ is above the level of concern, the higher tier risk assessment must be performed either through a refinement of the exposure side of the equation, or through the effects side of the equation with higher tier studies such as microcosm or mesocosm tests.

Where such studies are requested, assessors should understand the relevance of different end-points associated with such tests, and how to properly evaluate the results.

8.9 INTERPRETATION OF RISK QUOTIENTS - LEVELS OF CONCERN

The risk quotient is interpreted through comparison with levels of concern (LOC) to analyse potential risk to non-target organisms. This comparison also informs the consideration of further testing or refinement, or regulatory action. Implicitly built into these LOCs are assessment factors. For example, an acute LOC of 0.1 means the PEC/(LC/EC50) has an assessment factor of 10 built into the effects value.

There is no international agreement on the levels of concern used to justify further work or regulatory action. Australia's LOCs are provided below for non-target organisms and compared to those used in the US and EU. It should be noted that the EU uses a toxicity-to-exposure ratio (TER) approach that is the inverse of the exposure-to-toxicity RQ approach used in Australia and the US (except in the case of bees and non-target arthropods where a hazard quotient is used). Consequently, the TERs used in the EU have been converted to RQ values for this exercise. Risk is deemed satisfactory where RQ values are below the levels of concern in the following tables.

Chemicals that are found to be persistent, bioaccumulative and toxic (Chapter 7) are not always amenable to an RQ approach and must be handled on a case-by-case basis in consultation with the regulatory agencies. Unless indicated otherwise, the information on EU values has been obtained from Annex VI to Council Directive 91/414/EEC and the US EPA values have been obtained from guidance on <http://www.epa.gov/oppefed1/ecorisk_ders/toera_risk.html>.

Table 8-1: Levels of concern for birds and mammals

Risk presumption	Risk quotient – levels of concern[1]			Equivalent assessment factors		
	EU[1]	US EPA	**Australia**	EU	US EPA	**Australia**
Acute risk	0.1[2]	0.5[3]	**0.1**	10	2	**10**
Acute restricted use	-	0.2[4]	**-**	-	5	**-**
Acute endangered risk	-	0.1[5]	**-**	-	10	**-**
Short term risk	0.1[2]	-	**-**	10	-	**-**
Chronic risk	0.2[2]	1[6]	**1**	5	1	**1**

1) Apart from toxicity to predict risk, the EU also uses the BCF to birds and other non-target terrestrial invertebrates, and authorisation will not be granted where the bioconcentration factor (BCF, related to fat tissue) is greater than 1, unless it is clearly established through an appropriate risk assessment that under field conditions no unacceptable effects occur – directly or indirectly - after use of the plant protection product according to the proposed conditions of use.
2) For levels of concern to be exceeded, it must be clearly established through an appropriate risk assessment that under field conditions no unacceptable impact occurs after use of the plant protection product according to the proposed conditions of use.
3) Calculated based on PEC/LC_{50} or LD_{50}/ft^2 or LD_{50}/day.
4) Calculated based on PEC/LC_{50} or LD_{50}/ft^2 or LD_{50}/day or LD_{50}<50 mg/kg.
5) Calculated based on PEC/LC_{50} or LD_{50}/ft^2 or LD_{50}/day.
6) Calculated based on PEC/NOEC.

Table 8-2: Levels of concern for terrestrial vertebrates, soil invertebrates, arthropods, and soil micro-organisms

Organism level	Risk Quotient – levels of concern			Equivalent assessment factors		
	EU	US EPA[1]	**Australia**	EU	US EPA	**Australia**
Honeybees	50[2]	-	**1**	0.02	-	**1**
Beneficial arthropods	2[3]	-		0.5	-	
Earthworms						
Acute	0.1	-	**0.1**	10	-	**10**
Chronic	5	-	**1**	2	-	**1**
Soil micro-organisms[4]	See below	-	**See below**	-	-	**-**

1. Currently, the US EPA does not characterise residue exposure for honey bees and other beneficial insects, hence no risk quotient can be determined.
2. The quotient may be >50 if it can be clearly demonstrated that under field conditions there are no unacceptable effects on honeybee larvae, behaviour or colony survival and development after use of the pesticide (Annex VI, Council Directive 91/414/EEC).
3. The quotient value of 2 given above is based on reasoning from ESCORT 2. Annex VI to 91/414/EEC states that where exposure will occur to non-target arthropods, no authorisation will be granted if more than 30% of the test organisms are affected in lethal or sub-lethal laboratory tests conducted at the maximum proposed application rate, unless it is clearly established that under field conditions there is no unacceptable impact on those organisms after use of the pesticide. Any claims for selectivity and proposals for use in integrated pest management systems shall be substantiated by appropriate data.
4. In the case of non-target soil micro-organism exposure, an unacceptable risk is presumed in the event that nitrogen turnover or carbon mineralisation is affected by more than 25% (after 100 days in the EU, or in the relevant study period in the case of Australia). To mitigate this risk, it must be demonstrated there is no unacceptable impact on microbial activity under field conditions, taking account of the ability of micro-organisms to multiply.

Table 8-3: Levels of concern for terrestrial plants

Risk presumption	Risk Quotient – levels of concern			Equivalent assessment factors		
	EU	US EPA	**Australia**	EU	US EPA	**Australia**
Acute High risk	0.2[1]	1[2]	**0.1[4]**	5	1	**10**
Acute endangered risk	-	1[3]	**-**	-	1	**-**

1. Described in EC 2002 (b). Effects are based on the ER_{50} value. If the TER based on the most sensitive species is greater than 5 (RQ <0.2) then effects on non-target plants are considered acceptable. This trigger of 5 presupposes that at least six species have been tested. The trigger may be reduced if information on more species is available.
2. Based on PEC/EC_{25}.
3. Based on PEC/(EC_{05} or NOEC).
4. The LOC = 0.1 is the case where an LC50 has been used. More flexibility can be applied where an EC_X value is used depending on circumstances. For example, an EC25 based on plant fresh weight may allow use of an LOC = 0.5.

Table 8-4: Levels of concern for aquatic organisms (including sediment organisms)

Risk presumption	Risk quotient – levels of concern			Equivalent assessment factors		
	EU[1]	US EPA	**Australia**	EU	US EPA	**Australia**
Acute High risk	0.01	0.5[2]	**0.1**	100	2	**10**
Acute restricted use	-	0.1[2]	**-**	-	10	**-**
Acute endangered risk	-	0.05[2]	**-**	-	20	**-**
Chronic risk	0.1	1[3]	**1**	10	1	**1**

1. Apart from toxicity of compounds in determining risk, the EU also uses BCF values and where the maximum bioconcentration factor (BCF) is greater than 1000 for plant protection products containing active substances which are readily biodegradable or greater than 100 for those which are not readily biodegradable, authorisation is not granted unless it is clearly established through an appropriate risk assessment that under field conditions no unacceptable impact on the viability of exposed species (predators) occurs - directly or indirectly - after use of the plant protection product according to the proposed conditions of use.
2. Based on PEC/(LC/EC_{50}).
3. Based on PEC/NOEC.

Table 8-5: Levels of concern for aquatic plants

Risk presumption	Risk quotient – levels of concern			Equivalent assessment factors		
	EU	US EPA	**Australia**	EU	US EPA	**Australia**
Acute High risk	0.01[1]	1[2]	0.1	100	1	10
Acute endangered risk	-	1[3]	-	-	1	-
Chronic risk	-	-	1	-	-	1

1) based on the algal growth inhibition effects value.
2) Based on PEC/EC_{25}
3) Based on PEC/(EEC_{05} or NOEC).

8.10 RISK MANAGEMENT OPTIONS

In many countries, there is no distinction between risk assessment and risk management. This is not the case in the US and Canada, where the risk manager weighs the risks in light of potential benefits. In Australia, risk assessment is generally considered separately from risk mitigation, but both are used when doing a risk assessment. To a certain degree, a risk assessment is an iterative process, whereby both refinements to the risk characterisation as well as options for risk mitigation are explored if Tier I results indicate an unacceptable risk.

While many of the following options have been used as risk mitigation alternatives in Australian assessments in the past, the following discussions are paraphrased from the relevant EU documents: EC, 2002 (a), (b) and (c):

8.10.1 BIRDS AND MAMMALS

If risk is deemed unacceptable on the basis of the RQ, then risk mitigation strategies must be considered. Risk mitigation strategies for birds and mammals generally include changes to application techniques or the timing of application. Such options are generally specific to the use of the product. Several options for spray applications, seed treatments, granules and various taxa are discussed below.

In addition, the most recent guidance from the EU with respect to risk assessment to birds and mammals raises risk refinement through metabolism and avoidance with application of body-burden models and dietary toxicity data (EFSA, 2008). It appears the actual guidance in this area is still to be finalized.

8.10.1.1 Risk from spray application

If a risk to birds and mammals has been indicated from the use of a spray, then a reduction in the application rate and/or application frequency may lower the risk. However, this may significantly affect the efficacy of the product. Alternatively, spot or row treatment may be appropriate depending upon the pest or disease being treated. While changing the method of application from a spray to a more targeted approach (eg bait or paste/paint) may reduce the risk to birds and mammals, the success of this approach will depend upon the disease or pest being treated. If a reproductive risk to birds or mammals has been highlighted, then it may be appropriate to restrict the time of application to when birds or mammals are not breeding, or to limit the number of applications and hence reduce exposure.

8.10.1.2 Risk from seed treatments

If a high risk from a seed treatment is predicted then a label instruction should require the immediate removal of spills. Furthermore, it may be appropriate to consider that the seed be drilled or incorporated immediately after application. Availability of seed to birds and mammals will be reduced if seed is incorporated at sowing. Consequently, acute risk will be reduced because birds and mammals will take longer to find and consume treated seed. Nevertheless, consideration must be given to whether consumption will be sufficiently reduced that the risk is acceptable. In considering such an option, agronomic practices should also be assessed. For example: Will the seed still germinate? Will the seed treatment be effective if the seed is incorporated? This risk management option has been considered in detail by Pascual *et al*, 1999(a) and further information regarding risk management options for cereal seed is presented in Pascual *et al*, 1999(a) and (b).

8.10.1.3 Risk from granules

If a high risk from granules has been highlighted, again removal of spills should be required. Again, the feasibility of incorporating them at the time of application should be considered in order to reduce the availability to birds. As with seed treatment, agronomic implications should be considered when assessing this as a risk management option.

8.10.1.4 Risk from rodenticides

The availability of bait to non-target birds and mammals can be reduced by prescribing burrow-baiting or the use of bait stations. When surface spreading is necessary, then application should be on vegetation rather than on bare soil. Removal of dead and moribund rodents as well as removal of bait remains after completion of the control operation should be regarded as routine safety measures, particularly for rodent control in and around buildings.

The practicality of whatever risk management option is chosen should be assessed to ensure that it does not reduce the effectiveness of the product.

8.10.2 BEES

The risk mitigation measures outlined below for bees are only options. Such options will heavily depend on local agronomic practice and conditions. If predicted effects to honeybees are considered not acceptable, the following aspects of the use pattern may be considered for modification in order to mitigate the predicted risk:

- application rate
- timing of application (eg apply in the evening after honeybee flight, do not apply during honeybee flight)
- adaptation of agricultural practices (eg do not apply during crop flowering)
- agronomic practice (eg mulch ground cover before application of the plant protection products).

The APVMA and DEWHA have negotiated rewording of bee advice for labels through the RLC's Labelling Code Working Group. As well as agreeing that statements previously in use were impractical, it was also agreed that it is inappropriate for DEWHA to provide advice on livestock management, and that the label statement it proposes should indicate the nature of the hazard to bees. A recent example is:

> "Dangerous to bees. Will kill bees foraging in the crop to be treated or in hives which are over-sprayed or reached by spray drift. Residues may remain toxic to bees for several days after application."

Such label statements allow the user to determine an appropriate risk management strategy. The agreed label statements were subject to consultation around the various state and territory Departments of Primary Industry in 2006. There needs to be some flexibility in the application of such statements depending on the nature of the active constituent and the use situation.

8.10.3 OTHER ARTHROPODS

The following use specifications may be modified in order to reduce effects on non-target arthropods within the cropped area:

- application frequency and intervals
- timing of application (crop stage)
- unsprayed headlands.

The following options may reduce effects in off-field areas:

- buffer zones
- wind breaks
- drift-reducing application techniques.

For further explanations refer ESCORT 2 (Candolfi *et al*, 2001).

8.10.4 SOIL MICRO-ORGANISMS

Risk mitigation options for soil organisms are extremely limited. There are possibilities to reduce the exposure (reduction of application rate and/or number of applications and/or restriction on glasshouse use only), but inevitably these measures will compromise the agricultural objectives.

8.10.5 NON-TARGET PLANTS

Risk mitigation options for risks to non-target plants are similar to those for non-target arthropods in off-field areas, as follows:

- buffer zones to sensitive areas
- drift-reducing application techniques in the vicinity of sensitive areas.

8.10.6 AQUATIC ORGANISMS

Where refinement of the exposure assessment (buffer zones, restrictions on application etc) still results in an unacceptable risk to aquatic organisms, higher tier toxicity testing may be considered. A separate discussion on this is found in Appendix VI.

CHAPTER 9 – VETERINARY MEDICINES

9.1 INTRODUCTION

Assessment of the potential environmental risks posed by veterinary medicines follows the same basic process as that for pesticides. First the assessor must identify and quantify release and exposure, then determine the effects that the product has. Exposure is compared to effects in order to determine risk. However, in contrast to assessments of pesticides, there is harmonised guidance available for determining the environmental impacts of veterinary medicines. This is detailed further in the following chapter.

Along with Canada, Australia is an observer in the VICH program, which is a trilateral (EU-Japan-USA) program aimed at harmonising technical requirements for veterinary product registration. Its full title is the International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products (VMP). VICH was officially launched in April 1996. The VICH website is found at: <http://www.vichsec.org/>.

Through this program, a guidance document on environmental impact assessment of veterinary medicinal products has been developed (VICH, 2004). Two phases of environmental assessment are recommended. In Phase I, the potential for environmental exposure is assessed based on the intended use of the VMP. It is assumed that VMPs with limited use and limited environmental exposure will have limited environmental effects and thus stop in Phase I. This is determined through the use of a decision tree. Phase I also identifies VMPs that require a more extensive environmental assessment under Phase II.

Some VMPs that might otherwise stop in Phase I may require additional environmental information to address particular concerns associated with their activity and use. These situations are expected to be the exception rather than the rule and some evidence in support of the concern should be available.

In an effort to harmonize the environmental assessment to the maximum extent possible, guidance is available for both Phase I and Phase II assessments. The assessment guidance for Phase I assessments is available at <http://www.vichsec.org/pdf/2000/Gl06_st7.pdf>.

Phase II guidance is available at <http://www.vichsec.org/pdf/10_2004/GL38_st7.pdf>. The document provides a common basis for environmental impact assessments for veterinary medicinal products (VMPs) between the EU, Japan, US, Canada and Australia/New Zealand. It is recognised that significant regional differences (eg animal husbandry practices, climates, soil and water types, etc) preclude fully harmonised guidance at this time. Full harmonisation on principles of fate, effects and risk assessment is possible. However, the parameterisation and decision making is the prerogative of the individual regulatory authority. For this reason, the scope and extent of information recommended for environmental assessments for all regions cannot be completely specified. To the extent possible, the above document provides recommendations for standard datasets, and conditions for determining whether more information should be generated for a given VMP.

The data requirements stipulated through VICH will be encompassed within the revised Australian data requirements outlined in Chapter 3. Assessors should consult the VICH data requirements for specific data elements needed to perform the assessments.

The general elements of the framework are reproduced in this chapter, including:

- Protection goals (Section 9.2)
- General description of the VICH (Section 9.3)
- Exposure of VMPs to the Environment (Section 9.4)
- Risk Quotient (RQ) Approach (Section 9.5)
- Test Guidelines (Section 9.6)
- Metabolites (Section 9.7), and
- Biodegredation data (Section 9.8).

9.2 PROTECTION GOALS

Legislation and policy on environmental quality in the VICH regions set out the protection goals reflected in the assessment. The overall target of the assessment is the protection of ecosystems.

The aim of the VICH guidance is to assess the potential for VMPs to affect non-target species in the environment, including both aquatic and terrestrial species. It is not possible to evaluate the effects of VMPs on every species in the environment that may be exposed to the VMP following its administration to the target

species. Consequently, the taxonomic levels tested are intended to serve as surrogates or indicators for the range of species present in the environment.

The impacts of greatest potential concern are usually those at community and ecosystem function levels, with the aim being to protect most species. However, there may be a need to distinguish between local and landscape effects. For example, there may be some instances where the impact of a VMP at a single location may be of significant concern, such as for endangered species or a species with key ecosystem functions. These issues should be handled through risk management at the specific location. Such risk management may include restriction or prohibition of use of the product of concern in that specific local area. Additionally, issues associated with cumulative impact of some VMPs may be appropriate at a landscape level. These types of issues cannot be harmonised although they need to be considered as part of the assessment.

9.3 GENERAL DESCRIPTION AND USE OF THE VICH GUIDANCE

The VICH guidance contains sections for each of the major branches: (1) aquaculture, (2) intensively reared terrestrial animals and (3) pasture animals, each containing decision trees pertaining to the branch. The document also contains a section that lists the recommended studies for physical/chemical properties, environmental fate and environmental effects, as well as a description of how to determine when studies may be relevant.

The guidance uses a two-tiered approach to the environmental risk assessment. The first tier, Tier A, makes use of simpler, less expensive studies to produce a conservative assessment of risk, based on exposure and effects in the environmental compartment of concern. If the assessment cannot be completed with such data - due to a prediction of unacceptable risk - then the applicant/sponsor progresses to Tier B.

In some cases, it may be possible to implement a risk management option instead of moving to Tier B. In these cases, discussion with the regulatory authority is necessary. It should be recognised that risk management may not be identical for all regions and where Tier B testing is omitted in one region, it may still be recommended in another.

For certain VMPs, it may be necessary to go beyond Tier B because more complex studies, specific to issues being addressed or to a particular region, are necessary to complete the risk assessment. Such studies cannot be comprehensively dealt with in a harmonised guidance document. Therefore, these issues do not fall within the purview of this document, but should be addressed on a case-by-case basis with the appropriate regulatory authority. Examples include exceeding relevant trigger values in Tier B, where further testing may be warranted and/or risk mitigation measures may need to be implemented. As risk management measures are not within the scope of this guidance document, no guidance on these aspects is possible.

9.4 EXPOSURE OF VMPS TO THE ENVIRONMENT

The route and quantity of a VMP entering the environment determines the risk assessment scenarios that are applicable as well as the extent of the risk assessment. This guidance sets out a number of emission scenarios, using various assumptions. Emission can occur at various stages in the life cycle of the product. There may be some emission scenarios that are not applicable to a specific region.

However, with the exception of certain topicals or those added directly to water, most VMPs first pass through the animal to which it is administered. Generally the most significant environmental exposure results from excretion of the active constituent, being the parent and/or its metabolites. Following excretion, residues are generally assumed as being uniformly distributed in the environment, even though distribution may be patchy.

9.5 RISK QUOTIENT (RQ) APPROACH

The assessment is based on the accepted principle that risk is a product of the exposure, fate and effects assessments of the VMP for the environmental compartments of concern. In this respect it mirrors the approach described in the rest of this manual for pesticides. The risk assessor should already be familiar with the tools and approach described in the rest of this manual. The assessment is based on a RQ approach, which is the ratio of the predicted environmental concentration (PEC) and the predicted no effect concentration (PNEC) on non-target organisms. The RQ (PEC/PNEC) is compared against a value of one. A value less than one indicates that no further testing is recommended. However, in some circumstances, professional judgement is needed for a final determination.

The PEC is defined as the concentration of the parent compound and metabolites predicted to be present in the soil, water and sediment compartment. Worldwide harmonisation of PEC calculations is not practical or possible at this time. Regional differences in animal husbandry practices, different environmental conditions in

the VICH regions, and differences in treatment rates and frequency should be taken into account when calculating PECs. Therefore, the guidance document does not contain any examples of PEC calculations but gives some general qualitative guidance needed to determine PECs. It is incumbent upon the applicant/sponsor to determine the most appropriate method of estimating exposures for the region of interest for a particular VMP, based on regulatory guidance.

The PNEC is determined from the experimentally determined effects end-point divided by an appropriate assessment factor (AF). The AF is intended to cover uncertainties such as intra- and inter-laboratory and species variation, the need to extrapolate from laboratory study results to the field, and from short-term to long-term toxicity (acute: chronic ratios). The value varies depending on the type of study conducted. As always, variation in the applied AF should be clearly justified in the submission.

AFs of between 1000 and 10 are generally used in the assessment. A factor of 1000 is designed to be conservative and protective and is applied when only limited data are available. This value may be progressively reduced to 10 as more evidence becomes available. Such evidence could include:

- availability of data from a wide variety of species, including those which are considered to represent the most sensitive species
- information from structurally similar compounds, to suggest that the acute to chronic ratio is likely to be lower than that for many other compounds
- information to suggest that the chemical is rapidly degraded and not repeatedly administered so as to lead to chronic exposure.

9.6 TEST GUIDELINES

The specific test guidelines/protocols recommended in VICH guidance document are those finalised by OECD/ISO. This has the advantage of ensuring that environmental studies are current and broadly acceptable to regulatory authorities on a worldwide basis. However, lack of a specific study recommendation does not eliminate the importance of data on the identified specific organism class. In these situations, it is up to the applicant/sponsor to seek guidance from the appropriate regulatory authority.

Finally, conducting studies in accordance with good laboratory practice (GLP) is a regional requirement. It is preferred that studies should be conducted using methods that allow for a data audit as may be necessary for some regions. It should be recognised that if studies are not conducted to GLP, they might not be accepted in some VICH regions.

9.7 METABOLITES

The fate of chemicals in the environment is dependent on their chemical/physical properties and degradability. These properties will vary between the parent compound and the individual excreted metabolites. For example, the latter may be more water-soluble than the parent compound and may also be more mobile and/or more persistent in the environment.

In general, the data will be generated on the parent compound, yet the risk assessment should also consider relevant metabolites. This is particularly so for pro-drugs that are efficiently metabolised into a single metabolite for which testing may be more appropriate.

Consideration of the excretion data is not initially recommended at Tier A, where a total residue approach should be taken and a PECinitial should be estimated. It should be assumed that the VMP is 100% excreted as the parent. If the RQ is ≥1 for one or more tested taxonomic levels, then metabolism/excretion data from the residues as well as the dossier requirements for adsorption, distribution, metabolism, excretion (ADME) should be considered as part of the PEC refinement.

Excreted metabolites representing 10% or more of the administered dose and which do not form part of biochemical pathways should be added to the active substance to allow the PEC to be recalculated. If the RQ is still ≥1 after PEC refinement and testing at Tier B, then guidance should be sought from the regulatory authority, including whether testing of the major environmentally relevant metabolites needs to be considered.

9.8 SPECIAL CONSIDERATION FOR BIODEGRADATION DATA

If the RQ is <1 for all taxonomic levels tested at Tier A then the assessment should normally cease. However, for persistent compounds (eg DT90 >1 year in soil based on an annual application), it may be necessary to recalculate the PECinitial due to the possibility of accumulation in the environment.

If there are specific concerns related to persistence and/or mobility, then the degradates formed during environmental fate studies may need further investigation. It should be noted that an individual substance may be both an excreted metabolite and a degradate in the environment. In both cases guidance should be sought from the regulatory authority.

CHAPTER 10 – PROBABILISTIC RISK ASSESSMENT

10.1 INTRODUCTION

Probabilistic risk assessment (PRA) is a developing tool in environmental risk assessments although it is not yet established in regulatory procedures in Australia and the EU. However, the USA does use PRA routinely in their ecological assessment process as described further in Section 10.4 below.

The goal of probabilistic environmental risk assessment is to estimate the likelihood and the extent of adverse effects toward species as a result of exposure to a substance. For this reason, probabilistic assessments can produce more meaningful outputs that quantify the type, magnitude and frequency of effects. They are generally based on the comparison of an exposure/environmental concentration distribution (ECD) with a species sensitivity distribution (SSD) derived from toxicity data (Verdonck *et al*, 2003). Calculation of such distributions requires adequate data. For example, calculation of a species sensitivity distribution requires data points such as NOECs from at least eight to 15 species, from eight taxonomic groups. For this reason, data requirements for conducting adequate probabilistic risk assessments can be quite high. Comment from the EC regarding data requirements for probabilistic methods is detailed further in Section 10.3. This section also lists some of the strengths and weaknesses of probabilistic methods, while guidance from the OECD is discussed in Section 10.2 immediately below.

10.2 OECD

There is no guidance at OECD level on undertaking probabilistic risk assessment of chemicals, although some probabilistic processes have been documented. The use of statistical extrapolation to determine PNECs is documented in Section 4.2 of the OECD's *Manual for Investigation of HPV Chemicals* (OECD, 2004 (a)). The text of this section is included below for information.

If a large data set from long-term tests for different taxonomic groups is available, then statistical extrapolation methods may be used to derive a PNEC. The main underlying assumptions of the statistical extrapolation methods are as follows (OECD, 1992):

- the distribution of species sensitivities follows a theoretical distribution function
- the group of species tested in the laboratory is a random sample of this distribution.

The effects assessment can be performed with a statistical extrapolation method if the database on species sensitivity distributions (SSDs) is sufficient for its application (Posthuma *et al*., 2001).

In general, long-term toxicity data are log-transformed and fitted according to the distribution function. A prescribed percentile of that distribution is used as a criterion. Several distribution functions have been proposed, including a log-triangular function, a log-logistic function, and a log-normal function. Aldenberg and Slob, 1993 refined the way to estimate the uncertainty of the 95th percentile by introducing confidence levels, and it usually results in calculation of the HC5 which is the estimated concentration that should protect 95% of species. The HC5 is considered to be equivalent to, or an estimation of, the *Maximum Tolerable Concentration* (MTC). This model was further refined by Aldenberg and Jaworska, 2000. If one increases the protection level to the 99th percentile, the uncertainty associated with estimating the protection level decreases.

An advantage of these methods is that they use the whole sensitivity distribution of species in an ecosystem to derive a PNEC instead of always taking the lowest long-term NOEC. However, such make several assumptions (such as that the distribution of the NOECs is symmetrical) which may limit their accuracy. As well, several other drawbacks to these approaches include:

- the lack of transparency by using this method compared to the standard approach
- the question of the representativity of the selected test species
- the comparability of end-points
- the arbitrary choice of a specific percentile
- a statistical confidence level.

DEWHA has used statistical extrapolation methods such as Species Sensitivity Distributions to complement other assessment methods in cases where there are sufficient data. In so doing, and when using a statistical extrapolation method to derive a PNEC, the following issues need to be considered in the assessment:

1. Clarification of the type of input data, that is, preferably reliable NOECs from chronic/long-term studies, full life-cycle or multigenerational studies.
2. Information on the mode of action of the substance that may help to identify and to evaluate the need to include possible sensitive taxonomic groups or to exclude possible overrepresentation of certain taxonomic groups.
3. The minimum species requirements, for example, representative species from the following taxonomic groups: fish, crustaceans, insects, algae, higher aquatic plants, and other groups not already represented. It is recognised that no internationally standardised test guidelines for long-term tests are currently available for some taxa mentioned above. This requirement can be adapted based on knowledge/reasoning regarding sensitive end-points and species, as well as knowledge on structure – activity and mode of action.
4. The minimum sample size (number of data). This issue is the subject of an ongoing debate. While OECD (1992) proposes a minimum of eight NOECs on species from different taxonomic groups, EC (2003 (a)) recommends 10 NOECs (and preferably more than 15) on species from eight taxonomic groups. Similar proposals have been made by Gibbons and Coleman, 2001 and de Bruijn *et al*, 1999.
5. How multiple data for one species are dealt with, for example, averaging comparable data, or selecting the most sensitive end-point when various data are available.
6. Statistical fitting procedures. That is the method must be mentioned and explained. The log-normal distribution is the preferred one for pragmatic reasons. In addition, a statistical method is to be used to test the goodness of fit. In addition to the Kolmogorov-Smirnov test, the Anderson–Darling goodness of fit test can be used as a criterion for the choice of a parametric distribution for data-rich data sets, because it gives more weight to the tails of the distribution. Results should be discussed in regards to the graphical representation of the species distribution. If the data do not fit any distribution, the left tail of the distribution (the lowest effect concentrations) should be analysed more carefully. Any choice of a specific distribution function should be clearly explained.
7. Estimated parameter. That is the concentration corresponding with the point in the species sensitivity distribution (SSD) profile below which 5% of the species occur, may be derived with a 50% confidence interval associated with this concentration (as an intermediate value in the determination of the PNEC).
8. Estimation of the PNEC, that is, the intermediate value may de divided by an appropriate assessment factor, if needed, to reflect any identified further uncertainties. If mesocosm studies are available, they should also be evaluated to decide on the assessment factor.

Deviations from these recommendations can be made on a case-by-case basis, through consideration of sensitive end-points, sensitive species, mode of toxic action and/or knowledge from structure activity considerations.

10.3 EUROPEAN UNION

Recently, the European Commission funded a workshop on the subject of probabilistic risk assessment for pesticides (Hart, 2001) that reviewed the state-of-the-science and made recommendations regarding implementation and research needs. The following discussion on probabilistic risk assessment is paraphrased from the report of this workshop unless otherwise indicated.

There is a widespread perception that probabilistic methods are data hungry. However, it is difficult to make a fair comparison between probabilistic and deterministic methods. Some existing examples of probabilistic assessments have increased the complexity of the mechanisms considered, as well as treating them probabilistically.

Many of the methods described as probabilistic require more data because they handle new parameters that are not used in current routine risk assessments (eg landscape features, life history data, non-standard routes of exposure and body burdens). If decision-makers or risk assessors need to include these parameters in higher-tier assessments, then more data will be required whether the methods are probabilistic or not.

It may be possible to use probabilistic methods in ways that do not require more data. Some probabilistic methods (eg probability bounds) are designed specifically for situations where data are very limited. Indeed, a number of participants considered that it is when data are limited that uncertainty is greatest and probabilistic methods are most needed. However, the usefulness of methods such as probability bounds for regulatory assessment of pesticides has yet to be evaluated.

Where extra data are required for a probabilistic assessment, some of these data (eg landscape features, life history data) are generic in nature, not pesticide-specific and therefore, only need to be generated once.

The main strengths of probabilistic risk assessment processes were also documented, noting these will only be realised if probabilistic methods are implemented and used in appropriate ways. The identified strengths were:

1. **Probabilistic assessments can produce more meaningful outputs** that quantify the type, magnitude and frequency of effects. They can express risk in terms of the probability that unacceptable effects will occur, and have more ecological meaning than other measures of risk (eg toxicity-exposure ratios). Therefore, they may contribute to providing a better basis for deciding what is ecologically acceptable. If communicated well, they should also be more meaningful to decision-makers and the public.
2. **Probabilistic methods can quantify variability.** They allow assessors to take account of the high level of natural variability that exists in both exposure and effects. This provides a more complete description of the full range of risks and avoids the problems associated with using worst-case assumptions (eg lack of consensus in defining the worst case, and the generation of unrealistically extreme assessments by combining multiple worst case assumptions). It could also facilitate recognition of regional variation in the assessment and decision-making process.
3. **Probabilistic methods can quantify uncertainty.** They encourage assessors to re-examine their assumptions and define them clearly. Thus, they encourage assessors to identify uncertainties and gaps in their knowledge. They enable uncertainties that can be quantified to be included when estimating the probability of effects. They therefore enable clearer (and potentially more quantitative) statements to be made regarding the confidence that can be placed in the assessment outputs. This should provide a better basis for decision-making.
4. **Probabilistic assessments can make better use of the available information**, such as using all available toxicity data to quantify variation between species, rather than using the toxicity to the most sensitive tested species.
5. **Probabilistic methods can help identify which factors have most influence on risk**, through the use of sensitivity analysis. This should improve the use of resources by enabling additional studies to be focused on key factors. It should also help in the development of risk mitigation measures, if they can be designed to manipulate the factors that have most influence on risk.
6. **Probabilistic assessment may be a useful alternative to field testing**, especially for organisms where field testing is very difficult (eg fish). When field studies are necessary, probabilistic assessment may help to identify the key questions and areas of uncertainty that the field studies should address.
7. **Probabilistic methods promote better science**. They encourage critical examination and improvement of risk assessment models. They encourage generation of appropriate data, by showing how it reduces uncertainty. They may facilitate the use in risk assessment of new approaches such as geographic information systems and population modeling. Also, they encourage interaction between different scientific disciplines.

Substantial weaknesses and threats associated with the use of probabilistic methods were also identified. Some of the weaknesses also apply to deterministic methods (eg difficulty of validation). The following table lists these weaknesses and threats along with identified actions that could be taken to address them:

Probabilistic methods are more complex than simple deterministic methods and may take longer to conduct. Learning to use them will require significant training time. They require special software and more ecological and statistical expertise, which may not be available to all stakeholders. Complexity could also cause confusion or lack of confidence in conclusions, and slow down or cause inconsistency in the regulatory process.	**Recommended action:** Develop guidance on general principles and specific methods, and a framework to show where they fit in the assessment process. Agree on terminology. Develop standard software at a level suitable for general EU use, and make it available via the internet. Develop case studies to serve as examples. Provide training including exchange visits and workshop invitations. Ensure the availability of expert advice and establish appropriate procedures for peer review.
At least some probabilistic methods require more data, and might lead to increased animal testing. Also, gaining access to existing data is difficult.	**Recommended action:** Develop case studies to test usefulness of probabilistic methods that can be used with limited data and incorporate expert judgement. Refine methods for estimating species sensitivity distributions from small datasets. Optimise the design and use of effects tests in order to minimise the use of animals. Develop a database of regional scenarios for use in risk assessment. Make existing data accessible. Use sensitivity analysis to target resources on key factors, and match the level of complexity to the needs of the assessment.
Probabilistic approaches and outputs are difficult to communicate, even though the outputs should be more meaningful. Poor communication would be a substantial barrier to their acceptance.	**Recommended action:** Work with stakeholders, non-governmental organisations and social scientists to develop effective methods of communicating risk estimates from probabilistic assessments. Develop harmonised terminology and guidance on how to report the methods and results of probabilistic assessments.
There is a risk of misleading results, which could arise in several ways. Inappropriate or imprecise data or assumptions will give unreliable results. Models that are designed inappropriately, over-standardised, oversimplified, or exclude important factors, will also give misleading results. The complexity of the methods could lead to mistakes, provide the opportunity for abuse, give a false sense of precision, or cause different users to obtain different results from the same data. Inappropriate or incomplete quantification of uncertainty could also give a false sense of precision. These difficulties lead to concerns that probabilistic assessments might underestimate risk.	**Recommended action:** Establish standard procedures for explicit reporting of assessments, including all the assumptions that are made, to enable critical evaluation of all stages. Develop guidance on best practice for use in evaluating assessments. Use probabilistic assessments together with other lines of evidence in decision-making. Establish appropriate procedures to peer-review probabilistic assessments.
There is no established guidance on what outputs are required from probabilistic assessment. This makes it difficult to design an appropriate assessment. **There are no established criteria for interpreting the outputs of probabilistic assessment.** This would lead to delays, conflict and inconsistencies in decision making.	**Recommended action:** Begin every assessment by clearly defining what the output will be and the purpose or protection goal. In the short term, develop case studies for a workshop with stakeholders and decision-makers, and seek agreement on what the outputs should be and how to interpret them. In the longer term, carry out research to link individual effects to population and community effects and use the results to develop more ecologically appropriate regulatory triggers and protection goals.
Validation of probabilistic methods is difficult, and in the strict sense may not be completely achievable. The same is true for existing assessment methods.	**Recommended action:** Identify parts of the process that it will be possible to validate, calibrate or verify and implement research programs to do this. Compare results of probabilistic assessments to other evidence (eg field studies, incident monitoring).

This workshop concluded that the development of soundly based, well-presented guidance is an essential requirement for the successful implementation of probabilistic methods. This guidance is required both to assist people conducting probabilistic assessments, and to assist people who have to evaluate and interpret probabilistic assessments or use them in decision-making.

The workshop identified a need for several types of guidance:

- guidance on general principles
- guidance on specific methods
- guidance for standardised reporting of assessments, designed to enable critical evaluation of all stages
- the question of the representativity of the selected test species
- harmonised terminology.

The need for guidance on evaluating assessments is especially urgent, as probabilistic assessments are already being submitted to regulatory authorities but guidance has not yet been completed by the EU.

10.4 UNITED STATES

As noted above, the USA uses PRA routinely in their ecological assessment process. This follows from an initiative in 1996, where the US EPA presented two ecological risk assessment case studies to the Scientific Advisory Panel (SAP) for review and comment. While recognising and generally reaffirming the utility of US EPA's then-current deterministic assessment process, the SAP offered a number of suggestions for improvement. Foremost among their suggestions was a recommendation to move beyond the existing single point assessment process by developing the tools and methodologies necessary to conduct a probabilistic assessment of effects. Such an assessment would estimate the magnitude and probability of the expected impact and define the level of certainty and variation involved in the estimate. Following these recommendations, the agency began an initiative to refine the ecological risk assessment process for pesticides.

The key goals and objectives of EPA's initiative include the following:

- develop a conceptual approach to refine the ecological assessment process
- incorporate probabilistic tools and methods to provide an estimate of the magnitude and probability of effects
- address the broad spectrum of responses to pesticide exposure
- establish more realistic actual use scenarios and field conditions
- build upon existing data requirements for registration
- utilise, wherever possible, existing data bases and create new ones from existing data sources to minimise the need to generate additional data
- focus additional data requirements on reducing uncertainty in key areas.

The initiative began with the formation of The Ecological Committee on FIFRA Risk Assessment Methods (ECOFRAM), a stakeholder workgroup that was tasked with identifying and developing probabilistic tools and methods for terrestrial and aquatic assessments under the FIFRA regulatory framework. The conclusions and recommendations of the ECOFRAM were summarised in the draft reports, which were peer reviewed during two public workshops.

Once the reports and the peer review workshops were completed, the US EPA formed the Refined Risk Assessment Implementation Team (implementation team), that was charged with developing a plan to incorporate probabilistic tools and methods into the assessment process. After evaluating the ECOFRAM reports and workshop comments, the implementation team developed a conceptual approach for implementing changes to the then-current deterministic assessment process, using the reports and workshop comments as a starting point. This approach, which was evaluated and endorsed by the SAP in 2000, is based on a four-level risk assessment scheme. The levels of refinement within this scheme are as follows:

> **Level I:** This level most closely corresponds with current Australian screening approaches described in this manual. The conceptual risk assessment process for both aquatic and terrestrial assessments begins with Level 1, in which effects and exposure data are integrated to evaluate the potential for adverse ecological effects to non-target species. Level I provides a conservative screening level assessment based on the calculation of a risk quotient in which a point estimate of exposure is divided by a point estimate of effects. The magnitude and probability of risk are not evaluated in a Level I assessment. In this assessment, the PEC - based on maximum application rates and/or rates associated with other label options such as typical uses - is compared to an effects level, such as an acute or chronic toxicity value.
>
> Once the risk quotient is calculated, it is compared to the US EPA's levels of concern (LOCs). These LOCs provide the US EPA with criteria to analyse potential risk to non- target organisms and to consider the need for regulatory action.
>
> **Level II:** This level provides an initial estimate of the probability and magnitude of effects in vulnerable areas. Although this level provides point estimates for some parameters where little or no data are available for generating probability distributions, reasonable hypothetical distributions of exposure and effects parameters may be established using expert judgement and available published data. These distributions may be largely generic and are not necessarily species- or pesticide-use specific. Examples include distributions of residues on avian food items and metabolism of pesticides within and between soil

and water. Through sensitivity analysis, Level II assessments will identify the parameters that provide the greatest contribution to the variability and uncertainty of the assessment's conclusions.

Level III: Level III assessments will provide more refined predictions of the probability and magnitude of impacts. They will focus on exposure and effects parameters that contribute the most to the risk assessment uncertainty identified in the sensitivity analysis of the Level II assessment.

Level IV: Assessments at this level will provide the highest level of refinement. These assessments may include highly specific pesticide use scenarios and may incorporate additional data to establish the spatial and temporal pattern of exposure for species of concern. Additionally, data may be required to reduce the uncertainty associated with using effects data generated in laboratories for test species other than the focal species of concern. These data may include laboratory testing of the focal species themselves and effects testing conducted under actual field conditions of pesticide use.

The US EPA processes for probabilistic risk assessment of pesticides may be downloaded from the following site:<http://www.epa.gov/superfund/programs/risk/rags3adt/>.While Canada does not have their process readily documented, it appears they follow very similar processes to the US EPA in terms of tiers of assessment and the use of probabilistic approaches <http://www.pmra-arla.gc.ca/english/pdf/pmac/2002/pmac_25112002_d-e.pdf>. As well, data requirements between the two countries are generally harmonised.

10.5 AUSTRALIA

Australia does not have a formal position on the use of probabilistic tools in risk assessments. On occasions, PRA has been used in lieu of the deterministic screening approach for existing pesticides that are very data rich. However, datasets of this size are not common. Consideration of further work on this issue could potentially be pursued within the context of NChEM.

A combination of statistical distribution and assessment factor methods was established to derive toxicant guidelines in the 2000 edition of the Australian and New Zealand Guidelines for Fresh and Marine Water Quality, which can be downloaded at <http://www.environment.gov.au/water/publications/quality/index.html#nwqmsguidelines>.

GLOSSARY

ac	Active constituent
Acute toxicity	The ability of a substance to cause an adverse effect soon after a single exposure or dose. Any adverse effect resulting from a single short-term exposure to a substance
AF	Assessment factor
APVMA	Australian Pesticides and Veterinary Medicines Authority
BAF	Bioaccumulation factor
BCF	Bioconcentration factor
BMF	Biomagnification factor
BOD	Biochemical oxygen demand
Chronic Toxicity	The capacity of a substance to cause long-term adverse effects
CMR	Carcinogenic, mutagenic and toxic to reproduction
DOC	Dissolved organic carbon
DTx	Time taken for x% of a substance to be removed from a system
ELS	Early life stage
EPPO	European Plant Protection Organization
EqP	Equilibrium partitioning
EST	Emission Scenario Document
EUSES	European Union System for the Evaluation of Substances
FLC	Fish life cycle
Fugacity	A measure of the tendency of a substance to move from one phase to another or from one site to another
GHS	Globally Harmonised System of Classification and Labelling of Chemicals
GLP	Good laboratory practice
HQ	Hazard quotient
in vitro	In an experimental situation outside the organism. Biological or chemical work done in the test tube rather than in living systems
in vivo	In a living cell or organism
IOBC	International Organization for Biological and Integrated Control of Noxious Animals and Plants
ISO	International Standards Organization
L(E)C_x	The concentration of a substance that will be lethal (L) or induce an effect (E) to x% of the test population

LOAEC	Lowest observed adverse effect concentration
LOEC	Lowest observed effect concentration
MAF	Multiple application factor
MSDS	Material safety data sheet
NOAEC	No observed adverse effect concentration
NOEC	No observed effect concentration
OECD	Organisation for Economic Cooperation and Development
OPPTS	(US EPA) Office of Prevention, Pesticides and Toxic Substances
PBT	Persistent, bioaccumulative and toxic
PEC	Predicted environmental concentration
PNEC	Predicted no effect concentration
POP	Persistent organic pollutant
PRA	Probabilistic risk assessment
QSAR	Quantitative structure activity relationship
RQ	Risk quotient
Secondary poisoning	The poisoning of a predator or scavenger that eats a poisoned organism
SMILES	(**S**implified **M**olecular **I**nput **L**ine **E**ntry **S**ystem) string - a linear notation for chemical structures
SSD	Species sensitivity distribution
TER	Toxicity Exposure Ratio
TG	Test guideline
ThOD	Theoretical oxygen demand
Trophic level	One of the hierarchical strata of a food web characterised by organisms that are the same number of steps removed from the primary producers
UVCB	Chemicals of unknown or variable composition, complex reaction products and biological materials
vB	Very bioaccumulative
VICH	International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products
VMP	Veterinary medicinal product
vP	Very persistent

REFERENCES

1. Aldenberg, T. and Jaworska, J., 2000: *Uncertainty of the hazardous concentration and fraction affected for normal species sensitivity distributions*. Ecotoxicology and Environmental Safety, 46, 1-18.
2. Aldenberg, T. and Slob, W., 1993: *Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data*. Ecotoxicology and Environmental Safety, 25: 48-63.
3. APVMA, 2008. APVMA Operating Principles in Relation to Spray Drift Risk. Australian Pesticides and Veterinary Medicines Authority (APVMA), 15 July 2008.
4. Barrett K.L., Grandy N., Harrison E.G., Hassan S. and Oomen P. (eds.), 1994: *Guidance Document on Regulatory Testing Procedures for Pesticides with Non-Target Arthropods*. SETAC Europe, Brussels. ISBN 0 9522535 2 6.
5. Birkved M and Hauschild M. PESTLCI - A Pesticide Distribution Model for LCA. *Development of a Pesticide Distribution Model for use in Lifecycle Inventory Analysis*. Institute for Product Development. Technical University of Denmark. Lyngby, April 2003.
6. Boesten J, Helweg A, Businelli M, Bergstrom L, Schafer H, Delmas A, Kloskowski R, Walker A, Travis K, Smeets L, Jones R, Vanderbroech V, Van Der Linden A, Broerse S, Klein M, Layton R, Jacobson O and Yon D, 1997: *Soil Persistence Models and EU Registration*. The Final Report of the Work of the Soil Modelling Work Group of FOCUS (Forum for the Co-ordination of Pesticide Fate Models and their Use). 29 February, 1997.
7. Boller E, Vogt H, Ternes P and Malavolta C, 2005. Working Document on Selectivity of Pesticides (2005). International Organisation for Biological and Integrated Control of Noxious Animals and Plants.
8. Canada Gazette, 2000. *Persistence and Bioaccumulation Regulations. 2000*. Canada Gazette, Part II, Vol 134. No. 7., p 104.
9. Candolfi MP, Barrett KL, Campbell PJ, Forster R, Grandy N, Huet MC, Lewis G, Oomen PA, Schmuck R and Vogt H (eds) 2001: *Guidance document on regulatory testing and risk assessment procedures for plant protection products with non-target arthropods*. From the ESCORT 2 workshop. SETAC, Pensacola, 46 p.
10. Damstra T, Barlow S, Bergman A, Kavlock R and Van Der Kraak G (eds), 2002: *Global Assessment of the State-of-the-Science of Endocrine Disruptors*. An assessment prepared by an expert group on behalf of the World Health Organization, the International Labour Organization and the United Nations Environment Program. World Health Organization.
11. Davis BNK & CT Williams 1990: *Buffer zone widths for honeybees from ground and aerial spraying of insecticides*. Environmental Pollution (1990) 63:247-259. Published.
12. de Bruijn, J., Crommentuijn, T., van Leeuwen, K., van der Plassche, E. Sijm, D., van der Weiden, M. 1999: *Environmental Risk Limits in The Netherlands*, RIVM Report 601640001.
13. EC (European Communities), 1996. **Technical Guidance Document** in Support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and Commission Regulation (EC) No 1488/94 on Risk Assessment For Existing Substances. **Part II.**
14. EC (European Communities), 2003. **Technical Guidance Document on Risk Assessment** *in support of* Commission Directive 93/67/EEC on Risk Assessment for new notified substances; Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances; Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. **Part II.**
15. EC (European Communities), 2003 (b). **Technical Guidance Document on Risk Assessment** *in support of* Commission Directive 93/67/EEC on Risk Assessment for new notified substances; Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances; Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. **Part III.**
16. EC 2002 (a). Working document. *Guidance Document on Risk Assessment for Birds and Mammals Under Council Directive 91/414/EEC*. SANCO/4145/2000 – Final. 25 September 2002.
17. EC 2002 (b). DRAFT working document. *Guidance Document on Terrestrial Ecotoxicology Under Council Directive 91/414/EEC*. SANCO/10329/2002 ref 2 final. 17 October 2002.

18. EC 2002 (c). Working document. *Guidance Document on Aquatic Ecotoxicology in the context of the Directive 91/414/EEC.* Sanco/3268/2001 rev4 (final). 17 October 2002.

19. ECETOC, 2002. *Scientific Principles for Soil Hazard Assessment of Substances.* Technical Report No. 84. ISSN-0773-6347-84. Brussels, July 2002.

20. EFSA, 2008. Risk Assessment for Birds and Mammals. Revision of Guidance Ducument under Council Directive 91/414/EEC (SANCO/4145/2000 – final of 25 September 2002. European Food Safety Authority. The EFSA Journal (2008) 734: 1-181.

21. Environment Canada, 2003. *Guidance Manual for the Categorisation of Organic and Inorganic Substances on Canada's Domestic Substances List. Determining Persistence, Bioaccumulation Potential and Inherent Toxicity to Non-Human Organisms.* Existing Substances Branch, Environment Canada, June 2003.

22. EPPO (1994): *Decision-making scheme for the environmental risk assessment of plant protection products*, Chapter 11 Terrestrial vertebrates. EPPO Bull 24, 37-87.

23. EPPO (2001): *Guideline on test methods for evaluating the side-effects of plant protection products. No. PP 1/170(3). Honeybees.* EPPO Bull 31, 323-330.

24. EPPO (2002a): *Environmental risk assessment scheme for plant protection products.* Chapter 8. Soil organisms and functions. EPPO Bull, in prep.

25. Ganzelmeier, H., Rautmann, D., Spangenberg, R., Streloke, M., Herrmann, M., Wenzelburger, H-J. and Walter H-F. 1995: *Studies on the spray drift of plant protection products – Results of a test program carried out throughout the Federal Republic of Germany.* Mitteilungen aus der Biologschen Bundestalt fur Land- und Forstwirtschaft. Report 305. Blackwell Scientific, Berlin. Published.

26. Gibbons, R.D., Coleman, D.E. 2001: *Statistical methods for detection and quantification of environmental contamination,* John Wiley & Sons, New York, USA.

27. Hart ADM (ed.) 2001: *Probabilistic risk assessment for pesticides in Europe: Implementation & research needs. Report of the European Workshop on Probabilistic Risk Assessment for the Environmental Impacts of Plant Protection Products (EUPRA).* Published by Central Science Laboratory, Sand Hutton, York, UK, 109 pp.

28. Jager T, den Hollander H, Janssen G, van der Poel P, Rikken M and Vermeire T, 2000. *Probabilistic Risk Assessment for New and Existing Chemicals: Sample calculations.* RIVM Report 679102 049. National Institute of Public Health and the Environment, The Netherlands.

29. Kelly B C, Ikonomou M G, Blair J D, Morin A E and Gobas F A P C, (2007). Food Web–Specific Biomagnification of Persistent Organic Pollutants, Science:Vol. 317. No. 5835, pp. 236 – 239.

30. Lee-Steere C, 2005*: Environmental Risk Assessment Manual For the Proposed National Environmental Risk Management Framework for Chemicals For the Environmental Risk Assessment of: Industrial Chemicals, Pharmaceuticals, Food Additives and Non-Agricultural Biocides.* Draft Report Prepared for Department of Environment and Heritage.

31. Linders J, Mensink H, Stephenson G, Wauchope D and Racke K, 2000: *Foliar Interception and Retention Values After Pesticide Application. A Proposal for Standardized Values for Environmental Risk Assessment* (Technical Report. In *Pure Appl. Chem.,* Vol. 72, No. 11, pp 2199-2218.

32. Mensink B, de Greef, J and Linders J: *How to estimate concentrations of pesticides in surface water semi-quantitatively: an interim report* (in addition to EPPO bulletin 23, 1993). Report No. 679102011. National Institute of Public Health and the Environment/RIVM. Bilthoven, The Netherlands. December 1996.

33. OECD, 1992. *Report of the OECD Workshop on Quantitative Structure Activity Relationships (QSARs) in Aquatic Effects Assessment.* Environment Monograph No. 58.

34. OECD, 1994: *US EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships.* Environmental Monograph No. 88.

35. OECD, 1998: *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data.* OECD Series on Testing and Assessment, Number 10. Document ENV/MC/CHEM(98)18.

36. OECD, 2000: *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures.* Environmental Health and Safety Publications Series on Testing and Assessment No. 23.

37. OECD, 2001 (a): *Guidance Document on the Use of the Harmonised System for the Classification of Chemicals which are Hazardous for the Aquatic Environment.* Environmental Health and Safety Publications Series on Testing and Assessment No. 27.

38. OECD, 2001 (b): *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures.* OECD Series on Testing and Assessment, Number 33. 14 August, 2001, Paris.

39. OECD, 2002: *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals.* OECD Series on Testing and Assessment No. 21. ENV/JM/MONO(2002)8. 7 March 2002.

40. OECD, 2003 (a): *Draft Guidance Document for the Statistical Analysis of Ecotoxicity Data. OECD Environmental Health and Safety Publication.* Series on Testing and Assessment. Paris, 2003.

41. OECD, 2003 (b): *Persistent, Bioaccumulative and Toxic Pesticides in OECD Member Countries.* OECD Series On Pesticides Number 15. ENV/JM/MONO(2002)22. 12 February, 2003.

42. OECD, 2004 (a): *Manual for Investigation of HPV Chemicals*, OECD Secretariat, September 2004: <http://www.oecd.org/document/7/0,2340,en_2649_34379_1947463_1_1_1_1,00.html>

43. OECD, 2004 (b):. *Issues to be Addressed to Develop the Classification and Labelling for Terrestrial Environmental Hazards.* Task Force on Harmonisation of Classification and Labelling. ENV/JM/HCL(2004)3/REV. 13 July 2004.

44. Oomen PA, De Ruijter A and van der Steen J 1992: *Method for honeybee brood feeding test with insect growth-regulating insecticides.* EPPO Bull. 22, 613-616.

45. Pascual JA, Fryday SL and Hart ADM 1999(a): *Effects of Food Restriction on Food Avoidance and Risk of Acute Poisoning of Captive Feral Pigeons From Fonofos-treated Seeds.* Arch Env Contam Toxicol 37, 115-124.

46. Pascual JA, Hart ADM and Fryday SL 1999(b): *Incidence of Lethal Bird Poisoning Reduced by Regurgitation of Pesticide-Treated Food.* Environ Toxicol Chem 18, 247-253.

47. Pfleeger TG, Fong A, Hayes R, Ratsch H & Wiskliff C 1996: *Field evaluation of the EPA (Kenaga) nomogram, a method for estimating wildlife exposure to pesticide residues on plants.* Environmental Toxicology and Chemistry 15(4):535-543.

48. Posthuma, L., G.W. Suter II and T.P. Traas 2001: *Species sensitivity distributions in ecotoxicology.* CRR Press, Boca Raton Florida, USA.

49. Rautmann D, M Streloke & R Winkler 2001: *New basic drift values in the authorisation procedure for plant protection products.* In: Forster R & M Streloke eds, Workshop on risk assessment and risk mitigation measures in the context of the authorisation of plant protection products (WORMM), 27-29 September 1999. Mitt. Biol. Bundesanst. Land-Forstwirtsch, 133-141. Berlin-Dahlem, Heft 381.

50. Smith A, 1982: *Herbicides and the soil environment in Canada.* Canadian Journal of Soil Science. 62 (3): 433-460.

51. Streloke, M, Joermann G, Kula H, Spangenberg, R. 2002: *Analysis of toxicity data on aquatic organisms for regulatory purposes.* Poster at the SETAC-Europe meeting,Vienna.

52. UKAP, 2002: *Guidance Document on Regulatory Physico-Chemical Testing in the United Kingdon.* UK Analytical Partnership, July 2002.

53. United Nations, 2003: *Globally Harmonised System of Classification and Labelling of Chemicals (GHS).* <http://www.unece.org/trans/danger/publi/ghs/officialtext.html>

54. Urban DJ & Cook NJ 1986: *Hazard Evaluation Division Standard Evaluation Procedure: Ecological risk assessment.* EPA 540/9-85-001. Office of Pesticide Programs, United States Environmental Protection Agency, Washington, D.C.

55. US EPA, 1993: *Wildlife Exposure Factors Handbook.* Volume I of II. EPA/600/R-93/187. Office of Health and Environment Assessment. Office of Research and Development. December 1993.

56. US EPA, 1998: *Guidelines for Ecological Risk Assessment.* Federal Register 63(93):26846-26924. EPA/630/R-95/002F. April, 1998.

57. US EPA, 1998: *Guidelines for Ecological Risk Assessment.* Federal Register 63(93):26846-26924. EPA/630/R-95/002F. April, 1998.

58. US EPA, 1999: *Category for Persistent, Bioammumulative, and Toxic New Chemical Substances. Fe*deral Register: November 4, 1999. Volume 64, Number 213 pp 60194-60204. <http://www.epa.gov/fedrgstr/EPA-TOX/1999/November/Day-04/t28888.htm>

59. US EPA, 2000 (b): ECOSAR: available from <http://www.epa.gov/oppt/newchems/21ecosar.htm>

60. US EPA, 2000: *Pollution Prevention (P2) Framework*. Office of Pollution Prevention and Toxics, EPA-748-B-00-001. June, 2000.

61. US EPA, 2004: *Overview of the Ecological Risk Assessment Process in the Office of Pesticide Programs, US Environmental Protection Agency Endangered and Threatened Species Effects Determinations*. Office of Prevention, Pesticides and Toxic Substances Office of Pesticide Programs. Washington D.C., January 23, 2004.

62. Verdonck F, Aldenberg T, Jaworska J and Vanrolleghem P, 2003*: Limitations of Current Risk Characterisation Methods in Probabilistic Environmental Risk Assessment*. Environmental Toxicology and Chemistry, Vol. 22, No. 9.

63. VICH, 2004: *Environmental Impact Assessment for Veterinary Medicinal Products*. VICH GL 38 (Ecotoxicity Phase II). October, 2004.

64. Woods N and Dorr G, 2000:. *A consultancy to assess, measure and make recommendations on spray drift from the application of agricultural chemicals, Project 517 undertaken on behalf of Environment Australia*, The Centre for Pesticide Application & Safety (C-PAS), School of Agriculture and Horticulture, The University of Queensland, Gatton.

APPENDIX I – RELATIONSHIP BETWEEN VARIOUS PHYSICO-CHEMICAL DATA

Physico-chemical data are crucial because they are pivotal to both the risk assessment and the data requirements of the test package. In their guidance document on regulatory physico-chemical testing in the UK, the UK Analytical Partnership (UKAP) explains succinctly how various properties impact on other physico-chemical and ecotoxicology testing, as well as on the risk assessment. The following summary and discussion are taken from this document (UKAP, 2002):

Vapour pressure:

- extra care is needed to minimise vapour losses; related to boiling point
- impacts the choice of test method for biodegradation test
- impacts the choice of test vessels for ecotoxicity tests (eg closed systems to prevent vapour losses)
- important for the determination of atmospheric behaviour as for exposure of man via the environment calculations
- key parameter in determining environmental fate and behaviour leading to prediction of environmental concentrations.

Surface tension:

- not applicable for substances with a water solubility <1 mg/L
- may impact on the suitability of methods used for determining Kow and Koc for surface active substances
- will impact on the environmental fate of the chemical.

Water solubility:

- time to achieve saturation can be relevant to solution preparation for determining surface tension
- impacts on the concentration used in hydrolysis testing
- impacts on method for sample preparation for ecotoxicity tests
- important parameter leading to environmental classification and labelling
- key parameter in determining environmental fate and behaviour leading to prediction of environmental concentrations.

Partition coefficient:

- generally, substances with a high Log Kow will be hydrophobic and have low water solubilities and vice versa
- impacts the choice of test method for biodegradation test as some are not suitable for highly sorptive substances
- high Log Kow may lead to losses in ecotoxicity tests through adsorption
- important factor in determining bioaccumulation, adsorption potential and toxicity predictions
- used as a surrogate for bioaccumulation potential in the absence of these tests
- important parameter for environmental classification and labelling
- key parameter in determining environmental fate and behaviour leading to prediction of environmental concentrations.

The individual tests are related to each other in various ways. For example, water solubility and hydrolysis are closely linked. The solubility of the chemical is needed to select an appropriate test concentration for the hydrolysis study and in turn, knowledge of the hydrolytic behaviour of a substance is necessary to be able to interpret the results of a solubility study, since solubility is measured over a period of time.

Results of water solubility and hydrolysis should be available before determination of surface tension, partition coefficient (Kow) and adsorption coefficient (Koc) are conducted. The surface tension result can be used to judge if a Kow study is valid.

For hydrolysable substances, use of the HPLC methods for determination of Koc or Kow is advisable as they are faster than the wet-chemistry methods, although the HPLC methods may not apply for some substances such as metal complexes and surface-active substances. HPLC methods are more appropriate for substances that are poorly soluble in water and octanol. Results should be checked to ensure they are not conflicting (eg a highly water-soluble substance is unlikely to have a high Kow).

Similarly, values for melting point, boiling point and vapour pressure results should be checked for consistency. For example, a high melting-point solid is unlikely to have a high vapour pressure at ambient temperatures. Melting and boiling point results should also be considered when selecting the temperature range over which vapour pressure measurements are made, to ensure no phase transitions occurred during the determination.

APPENDIX II – INTERNATIONAL APPROACHES TO AQUATIC EXPOSURE ASSESSMENT

UNITED STATES

<http://www.eap.gov/oppefed1/models/water/geneec2_description.htm>

Development and Use of GENEEC Version 2.0 for Pesticide Aquatic Ecological Exposure Assessment May 1, 2001.

Background

The US EPA Office of Pesticide Programs (OPP) is required to assess the risk posed by pesticides to human health and the environment. The OPP Environmental Fate and Effects Division (EFED) is charged with carrying out the environmental portion of this assessment. To assess the risk to aquatic life posed by each chemical, EFED estimates pesticide concentrations that would be expected in the environment from normal use (exposure) and compares them to concentrations known to be toxic from laboratory tests (hazard). This exposure / hazard ratio is used as an indication of potential ecological risk to non-target species in the environment.

EFED has been performing pesticide aquatic exposure assessments as a part of the ecological risk assessment process for a number of years. Exposure assessments draw on both measured pesticide quantities in the field as well as computer modelling to establish the concentration levels that might be expected in significant aquatic habitats. With the development of enhanced environmental fate and transport models such as the Pesticide Root Zone Model (PRZM), (Carsel *et al*, 1984), Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) (Leonard *et al*, 1989) and the EXposure Analysis Modelling System (EXAMS) (Burns *et al*, 1991), computer modelling of pesticide exposure began to play a larger role in EFED's risk assessments in the early 1990's.

At that time, in response to EFED's need to have a standard aquatic environment in which all chemicals could be assessed and compared on an equal footing, a 'standard agricultural field-farm pond' scenario was selected for all aquatic exposure assessments. This standard pond scenario assumes that rainfall onto a treated, 10 hectare agricultural field causes pesticide-laden run-off into a one hectare; 20 000 cubic metre volume; 2 metre deep water-body. Although this 'standard scenario' was designed to predict pesticide concentrations in the standard farm pond, it has been shown to be a good predictor of upper level pesticide concentrations in small but ecologically important upland streams (Effland *et al*, 1999).

A system was developed using the electronically linked PRZM and EXAMS models and this standard pond scenario to simulate pesticide applications to most US agricultural crops simulating local soil, weather and farm management in the areas in which each are grown. Output from this modelling is daily pesticide concentrations in the standard farm pond over the 36 year period for which rainfall data is available. This became the EFED standard method for pesticide aquatic ecological exposure assessment.

As this method was relatively labour-intensive and therefore time consuming, a trigger or screening mechanism was used to establish which chemicals were most likely to pose higher risk and therefore should be assessed in this manner. The screening mechanism established was termed the 'back-of-the-envelope' calculation and was based largely on the solubility of the chemical. With the advent of better models together with doubts about the usefulness of a chemical's solubility as a relevant screen, work was begun in 1994 to develop a new screen that would be more consistent with other modelling approaches. It was also hoped that a new screen would also better represent the pesticide parameters that are linked to pesticide transport to, and persistence in, surface water. Accordingly, the GENEEC (GENeric Estimated Exposure Concentration) (Parker *et al*, 1995) model was developed, which replaced the back-of-the-envelope calculation in mid-1995 (World Wildlife Fund, 1992).

Development of a tiered system of modelling

Along with development of new modelling tools and methods, EFED has developed a tiered approach to determine the appropriate level of modelling needed to perform a risk assessment for each chemical. This tiered approach is designed to minimise the amount of analysis required to evaluate any given chemical. Each of the tiers is designed to screen out pesticides by requiring higher, more complex levels of investigation only for those that have not passed the previous tier. Each tier screens out a percentage of pesticides from having to undergo a more rigorous review prior to registration or re-registration. Passing a given assessment tier indicates that there is a low possibility of risk to the aquatic environment. Failing an assessment tier does not *a priori* mean the chemical is likely to cause environmental problems, but instead that the assessment should continue on to the next higher assessment tier. The end result of this tiered modelling system will ideally be as thorough an analysis

as is necessary for each pesticide. This enables the risk assessor and the regulator to focus greatest resources and efforts toward areas of greatest potential ecological threat. As a matter of policy, OPP does not take significant regulatory action based upon the results of tier 1 screening models.

Development of GENEEC Version 1.0

EFED had several criteria for development of a first tier screening model. First, the model should be fast and easy to use. Second, it should require only a few, readily available input parameters. Third, the input parameters should be those most significant to represent pesticide amount and type of application as well as transport to and persistence in surface water. Fourth, the predicted concentration values should be (1) higher than most of the highest of the values predicted in the next higher tier of modelling and (2) higher than most of the upper level concentration values that are measured in the field at vulnerable sites. The last requirement is designed to preclude the possibility that potentially hazardous chemicals pass the screen early in the assessment process and escape sufficient review. A vulnerable site is defined as one at which high concentration levels are expected due to the occurrence of those conditions of pesticide application, weather, and soils known to favour transport to and persistence in surface water.

GENEEC Version 1.0 was developed with these conditions in mind. It was designed to mimic a much more sophisticated PRZM/EXAMS simulation but requires far fewer inputs and much less time and effort to use. The model uses a chemical's label application information, its soil/water partition data and its degradation kinetics to estimate high level exposure values in the same EFED standard agricultural field/farm pond scenario as used with PRZM/EXAMS simulations. The program is generic in that it does not consider differences in climate, soils, topography or crop in estimating potential pesticide exposure.

GENEEC is also simpler than the PRZM and EXAMS models in its treatment of hydrology. The linked PRZM and EXAMS models simulate the impact of daily weather on the treated agricultural field over a period of 36 years. During this time, pesticide is washed-off of the field into the water-body by 20 to 40 rainfall/run-off events per year. Each new addition of pesticide to the water-body adds to the pesticide which has arrived earlier either through previous run-off events or through spray-drift. Moreover, the pesticide is assumed to begin degrading on the day it reaches the water. On the other hand, GENEEC is a single event model. It assumes one single large rainfall/run-off event occurs that removes a large quantity of pesticide from the field to the water all at one time. Longer-term, multiple-day average concentration values are calculated based on the peak day value and subsequent values considering degradation processes.

Differences between GENEEC Version 1.0 and Version 2.0

GENEEC Version 2.0 was developed in response to suggestions for improvements by model users, by the desire to stay current with the newer versions of the PRZM (Carsel, 1997) and EXAMS (Burns, 1997) programs upon which GENEEC is based and by availability of much improved data on spray drift and quantitative methods of estimation of offsite drift developed by the Spray Drift Task Force (SDTF). The main differences between versions 1.0 and 2.0 include:

(a) an entirely new binding curve to represent dissolved concentration as a function of Kd;

(b) the use of the binding parameter, Kd in preference to Koc to represent pesticide attachment to soil, to organic matter or to water-body bottom and suspended sediments;

(c) a change in the recommendation for depth of incorporation;

(d) a change in the timing of the single event rainstorm for chemicals which receive multiple applications;

(e) addition of a sub-routine from the SDTF to estimate spray drift; and

(f) a change in the time durations of the output values to better match the durations of relevant toxicity tests.

These changes were made as follows:

(a) Relationship between binding and the dissolved concentration

The main operator in GENEEC Version 1.0 is a pesticide's organic carbon normalised equilibrium partition coefficient (K_{oc}). The K_{oc} is defined as the equilibrium adsorption coefficient (K_d) normalised to the soil's organic carbon (OC) content. It is calculated by dividing the K_d value by the organic carbon fraction. Initial development of Version 1.0 of the program began by exploring the exposure impact of full range of chemical K_{oc} values on the dissolved pesticide concentration in the standard field/pond system using the linked PRZM1 and EXAMS2.94 models. This was accomplished by repetitively increasing the K_{oc} value in both programs by an order of magnitude, running each new simulation and then recording the resulting instantaneous dissolved concentration. This gave a series of dissolved pesticide concentrations in the pond

as a function of K_{oc}. The S-shape of the resulting curve suggested that a four-parameter Morgan-Mercer-Flodin 1975 type model might replicate the function most closely. This function was fit and the resulting curve was programmed into the model. The EXAMS parameter PRBEN remained at its default value of 0.5 in order to equally divide influent adsorbed pesticide between the water column and the bottom sediments.

The same process was followed in the development of Version 2.0, except that the K_d value was used in place of the K_{oc} while the PRZM3.12 and EXAMS2.97.7 models were used in place of the earlier versions of those programs. The curve described by the successive points was not easy to match with a standard fitting routine, so a process of linear interpolation between the points was used instead. The resulting relationship is used in both GENEEC Version 2.0 and FIRST Version 1.0.

Within the EXAMS program, the binding parameter (K_{oc} or K_d) controls not only the final equilibrium partitioning of the chemical between the dissolved (in water) and adsorbed (to soil or to bottom sediments) phases, but also the rate at which binding takes place to reach this equilibrium. For high K_{oc} or K_d values, the binding takes place largely within the first day. For lower values of K_{oc} or K_d, the process may not be complete for almost a year. In order to accurately mimic this EXAMS process in the GENEEC model, an empirical procedure was carried out to simulate a pseudo-binding rate as a function of K_{oc}. This rate was determined by turning off all degradation processes within the PRZM and EXAMS models and calculating an apparent rate constant that would account for the continuous decline in concentration values. Another series of PRZM-EXAMS simulations were run for K_{oc} values ranging from 10^{-1} to 10^{+8} and a daily pseudo-binding rate was determined. A four-parameter Morgan-Mercer-Flodin 1975 type function was then fit to calculate these apparent binding rate constants as a function of K_{oc}. This ongoing adsorption within the pond is then programmed to occur simultaneously with chemical and biological degradation processes. For these reasons, pesticide concentrations in Version 2.0 are likely to exceed concentrations in Version 1.2 by a small amount.

b) Use of the binding parameter, K_d in preference to K_{oc} to represent pesticide attachment to soil, to organic matter or to water-body bottom sediments

Adsorption (binding) tests are performed on soils of different textural classes, pH's and organic matter contents. The soil/water partition coefficient (K_d) is defined as the ratio between the concentration in soil and the concentration in water. Therefore, it can be used to estimate the dissolved or the adsorbed fraction in a soil-water system for all chemicals. For this reason it is the preferred parameter for this purpose and is recommended for use in GENEEC Version 2.0.

The organic carbon normalised soil/water equilibrium partition coefficient (K_{oc}) may be preferred for pesticides for which there is a strong positive correlation between the K_d value and the organic carbon content of the soils on which the adsorption tests were performed. If there is correlation, the multiple K_{oc} values will be less variable than the multiple K_d values and the K_{oc} is likely to be a more accurate estimator. If there is no correlation, use of the K_d is preferable. The K_d/K_{oc} conversion is based on an organic matter content of 2 percent and an organic carbon content of 1.16 percent. If neither the K_d nor the K_{oc} is available, use of 0.35 times the K_{ow} value is recommended (Burns, 2001, personal communication).

(c) Changes in recommendations for type and depth of incorporation

Version 1.0 of GENEEC recommended using a depth of 1.0 inch (2.54 cm) to represent in-furrow application. However, a literature review shows that a value of 2.0 (5.08 cm) would more accurately represent actual placement of the material. For a banded-incorporated application, version 1.0 recommended a depth of 2.0 inches. A depth of 1.2 inches (3.05 cm) would be a more accurate representation. Version 1.0 suggested that no incorporation was appropriate for either aerial or ground applications. However, Version 2.0 suggests no incorporation for aerial and airblast applications. Incorporation of ground applications would be variable depending on the type of application being made (eg. in-furrow ground spray incorporated to 2.0 inches).

(d) Change in the timing of the single event rainstorm for chemicals that receive multiple applications

Some pesticides are designed to degrade very quickly in the field so that in a short time, little residue remains. In order to give 'credit' to such pesticides, GENEEC was designed to allow a two-day degradation period for the pesticide in the treated agricultural field prior to the single rainfall event that washes the pesticide into the farm pond. The aerobic soil metabolism rate is used to simulate the decline during this two-day period as well as during the period between multiple applications prior to the rainfall/run-off event. The two-day period is not used if the pesticide label requires that the pesticide be 'wetted-in' at the time of application, either through irrigation or rainfall. In GENEEC Version 1.0, the two-day period between application and the rainfall event was provided only following a single application. For multiple applications,

the rainstorm was programmed to occur immediately on the premise that 'credit' for rapid degradation in the field had already been provided by rapid degradation between successive applications. However, this caused confusion when comparing the impact of single with multiple applications. Consequently, Version 2.0 is programmed to provide for the two-day degradation period in the field following either single or multiple applications.

(e) Addition of a sub-routine from the SDTF to estimate spray drift

In GENEEC Version 1.0, spray drift percentages and application efficiency factors were set to fixed values based on literature values and personal communication with experts in the field. Aerial and airblast spray drift directly to the pond was assumed to average 5% of the application rate across the 208 foot width of a one acre square pond. When the scenario was metricised from acres to hectares, the original width of the pond was kept. Ground spray drift directly to the pond was assumed to average 1% of the application rate across the width of the pond. Application efficiency was assumed to be 95% for aerial and airblast spray and 99% for ground spray.

GENEEC Version 2.0 incorporates much more sophisticated spray drift treatment based on availability of much improved data on spray drift and quantitative methods of estimation of offsite drift developed by the Spray Drift Task Force (SDTF). In Version 2.0, the user may simulate application by aerial spray, by airblast spray, by ground spray or by broadcast application of granular material. Spray drift is calculated using a sub-routine developed by SDTF for this purpose. It estimates the ninetieth percentile total down-wind deposition onto the two hundred and eight foot wide water body. The program allows the user to specify either aerial, airblast or ground application, select the spray quality (droplet size distribution) and simulate a no-spray zone between the treated field and the water (if one is required by the pesticide label). Note: Spray quality is the droplet size distribution as defined by American Society of Agricultural Engineers (ASAE) Standard 572 for all types of applications.

EFED default spray quality is specified for those cases where none is given on the pesticide label. The default is no buffer unless one is specified on the label. A safety factor of 3.0 is applied to the SDTF drift estimate for airblast application. In the case of broadcast application, 100% application efficiency is assumed with no pesticide drifting directly to the pond. Biological and chemical degradation of the spray drift in the pond is assumed to begin at the time the chemical enters the pond. Degradation of the pesticide reaching the pond via run-off begins on whichever day it reaches the pond. GENEEC calculates the contributions from spray drift and from run-off independently and then combines the results to give the final concentration.

(f) Change in the time durations of the output values to better match the durations of relevant toxicity tests

GENEEC Version 1.0 estimated the peak value which occurs on the day of the single large rainstorm as well as multiple day averages over periods of 4, 21, and 56 days. These averages were compared to the results toxicity tests carried out for the same durations in a screening level risk assessment.

GENEEC Version 2.0 estimates the peak value which occurs on the day of the single large rainstorm as well as multiple day averages over periods of 4, 21, 60 and 90 days. These averages are appropriate to comparison with the toxicity tests that are currently performed.

Other simulated processes

Overall degradation rate. Calculating degradation in the pond is for the purpose of estimating annual average concentration values for chronic exposure assessment. Degradation in the pond considers aerobic aquatic metabolism, abiotic hydrolysis and direct aquatic photolysis. The EXAMS program upon which the aquatic degradation is based considers the amount of light penetration in the simulated water body to calculate an effective photolysis rate based on the rate that is input into the model. Based on EXAMS simulation, the photolysis rate constant in the relatively murky pond is 124 times slower than that in clear water. The overall degradation rate in the pond is calculated by summing the aerobic aquatic metabolism rate constant, the abiotic hydrolysis rate constant and 1/124th of the direct aquatic photolysis rate constant.

Pesticide soil incorporation. The program also accounts for pesticide incorporation at the time of application. Incorporation reduces the mass of pesticide available to run-off by a factor equal to the depth of incorporation in inches up to a maximum of six inches.

Pesticide solubility. GENEEC considers the solubility of the chemical only as an upper limit on the dissolved concentration estimate. If the estimated concentration values exceed the user input solubility of the chemical, the concentration values are reduced to the solubility level.

Burns, L.A. 1991. *Exposure Analysis Modeling System: Users Guide for EXAMS II version 2.94*, Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA.

Burns, L.A. March 1997. Exposure Analysis Modeling System (EXAMSII) Users Guide for Version 2.97.5, Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA.

Carsel, R.F., C.N. Smith, L.A. Mulkey, J.D. Dean and P. Jowise. 1984. *Users manual for pesticide root zone model (PRZM)*: Release 1, Rep. EPA-600/3-84-109, 219 pp. Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA.

R.F. Carsel, J.C.Imhoff, P.R.Hummel, J.M.Cheplick and J.S.Donigian, Jr. 1997. PRZM-3, *A Model for Predicting Pesticide and Nitrogen Fate in Crop Root and Unsaturated Soil Zones: Users Manual for Release 3.0*; Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA.

Effland, W.R., Thurman, N.C., Kennedy, I. *Proposed Methods For Determining Watershed- Derived Percent Cropped Areas and Considerations for Applying Crop Area Adjustments To Surface Water Screening Models*; USEPA Office of Pesticide Programs; Presentation To FIFRA Science Advisory Panel, May 27, 1999.

Leonard, R.A., W.G. Knisel and D.A. Still. 1989. GLEAMS: *Groundwater loading effects of agricultural management systems.* Transactions American Society of Agricultural Engineering. 30: 1403-1418.

Morgan, P.H., L.P. Mercer and N.W. Flodin, 1975. *A General Model for Nutritional Responses of Higher Organisms.* Proceedings of National Academy of Sciences, USA; Vol.72 pp. 4327-4331.

Parker, R.D., R.D. Jones and H.P. Nelson., 1995. GENEEC: *A Screening Model for Pesticide Environmental Exposure Assessment.* In Proceedings of the International Exposure Symposium on Water Quality Modeling; American Society of Agricultural Engineers, pp. 485-490; Orlando, Florida.

World Wildlife Fund, *Improving Aquatic Risk Assessment under FIFRA*: Report of the Aquatic Effects Dialogue Group, pp 23-24, 1992.

EUROPEAN UNION

Available from model documentation <http://viso.ei.jrc.it/focus/sw/models/STEPS_ONE_TWO/download.html>

Introduction

As described in the remit of the surface water scenarios working group, step 1 and 2 calculations should represent worst-case loadings and loadings based on sequential application patterns, respectively. These loadings should not be specific to any climate, crop, topography or soil type. With this in mind the group developed two simple scenarios for calculating exposure in surface water and sediment and has constructed a MICROSOFT Visual Basic application for the derivation of PEC values in water and sediment.

The assumptions at both steps 1 and 2 are very conservative. The assumptions are essentially based around drift values calculated from BBA (2000) and an estimation of the potential loading of pesticides to surface water via run-off, erosion and/or drainage. This run-off loading represents any entry of pesticide from the treated field to the associated water body at the edge of the field.

At Step 1, inputs of spray drift, run-off, erosion and/or drainage are evaluated as a single loading to the water body and worst-case water and sediment concentrations are calculated. If inadequate safety margins are obtained (toxicity exposure ratios <trigger values), the registrant proceeds to Step 2. At Step 2, loadings are refined as a series of individual applications, resulting in drift to the water body, followed by a run-off/erosion/drainage event occurring four days after the last application. The amount lost via run-off is determined by the crop interception, the region of use (northern or southern Europe) and season of application. Again, if inadequate safety margins are obtained (toxicity exposure ratios <trigger values), the registrant then proceeds to Step 3. Step 3 requires the use of the deterministic models PRZM, MACRO and TOXSWA.

Please notice that the interception percentages used by steps 1-2 in FOCUS are not the same as listed in the FOCUS groundwater report (FOCUS 2000*, FOCUS groundwater scenarios in the EU plant protection product review process, Report of the FOCUS Groundwater Scenarios Workgroup*, EC Document Reference Sanco/321/2000, 197pp) because more recent literature has been used to compile the numbers.

The purpose of formalising Step 1 and Step 2 calculations is to harmonise the methods of calculation and to avoid unnecessary, complex exposure assessments for plant protection products when large safety margins exist even with conservative scenarios.

Standard assumptions

A set of assumptions for the water body dimensions that were common to step 1 and 2 were compiled to derive the scenario. These are based upon existing concepts with the EU and member states, together with expert judgement. These assumptions are as follows:

A water depth of 30 cm overlying sediment of 5 cm depth was selected in order to comply with existing risk assessment approaches within the EU and existing ecotoxicity testing requirements for sediment-dwelling organisms. The density of the sediment was selected to be 0.8 g.cm-3 with an organic carbon content of 5%. The water body is assumed to have an area equivalent to one tenth of the field from which it receives run-off or drainage water (a field:water ratio of 10). Assuming a 1 ha field, the 0.1 ha (1000 m2) water body will have a volume of 3 x 105 litres.

Step 1 assumptions

At Step 1, inputs of spray drift, run-off, erosion and/or drainage are evaluated as a single loading to the water body and worst-case surface water and sediment concentrations are calculated. The loading to surface water is based upon the number of applications multiplied by the maximum single use rate unless:

3 × DT50 in whole sediment/water systems < time between individual applications.

In such a case the maximum individual application is used to derive the maximum PEC as there is no potential for accumulation in the sediment/water system. All inputs are assumed to occur at the same time. However, their initial distribution between the surface water and sediment compartments is dependent upon the route of entry and sorption coefficient (Koc) of the compound. Drift inputs are loaded into the water where they are subsequently distributed between water and sediment according to the compound's Koc. The 'run-off' entry is distributed between water and sediment at the time of loading according to the compound's Koc and an effective sediment depth of 1 cm. In this way compounds of high Koc are added directly to the sediment whereas compounds of low Koc are added to the water column in the 'run-off' water.

Step 2 assumptions

At Step 2, inputs of spray drift, run-off, erosion and/or drainage are evaluated as a series of individual loadings comprising of drift events (number and timing as defined in Step 1) followed by a loading representing a run-off, erosion and/or drainage event four days after the final application. This assumption is similar to that developed by the US EPA in their GENEEC model. Degradation is assumed to follow first-order kinetics in soil, surface water and sediment (an option of using different degradation rates in surface water and sediment is included).

The fraction of each application reaching the adjacent water is both a function of method and number of applications. Drift values for aerial applications are not dependant upon the number of applications. Four days after the final application, a run-off loading is added to the surface water and associated sediment and is function of the residue remaining in soil (g/ha), region and season of application and the Koc/Kom.

The user selects from two regions (northern EU and southern EU according to the definitions given for crop residue zones and three seasons - February to May, June to September and October to January). If a product is used across both regions or two or more seasons then the Step 2 calculation can be used to evaluate the worst-case (according to the loadings defined in a look-up table) or to determine which combinations require further evaluation at Step 3.

The daily concentrations in surface water and sediment are calculated. The times of the maximum concentration in water and sediment and the actual concentrations 1, 2, 4, 7, 14, 21, 28, 42, 50 and 100 days after the peak in each phase (water and sediment) are reported. Then the time weighted average concentrations following the maximum concentration are calculated and reported for the same time periods. As in step 1, drift inputs are loaded into the water where they subsequently distributed between water and sediment according to the compound's Koc and an effective sediment depth of 1 cm. The run-off entry is distributed between water and sediment at the time of loading according to the compound's Koc. In this way compounds of high Koc are added directly to the sediment concentration whereas compounds of low Koc are added in the run-off water.

The fraction of the pesticide that enters the water body via drift has to be partitioned between water and sediment in the following days. As experimental data do not support a full partitioning within 24 hours, an extended approach is followed for STEPS1-2:

The pesticide is distributed in surface water into two theoretical compartments, available for sorption to sediment and unavailable for sorption to sediment according to the following equation:

masw = msw × K

musw = msw × (1-K)

where:

msw is total mass of pesticide in surface water (mg/m²)

masw is mass available for sorption and (mg/m²)

musw is mass unavailable for sorption (mg/m²)

K is the distribution coefficient (-), set to value of 2/3 for all compounds.

After the occurrence of the run-off/drainage event, it is assumed that full equilibrium between water and sediment is established within 24 hours (K = 1).

The multiple application pattern may lead to lower concentrations in surface water than the respective single application for certain combinations of application pattern and the degradation of the pesticide. Therefore, the program will always do a second run with the respective single application pattern if the user has entered a multiple application pattern.

Scenario data

Surface water definitions

Parameter:	Value
water depth (cm):	30
sediment depth (cm):	5
effective sediment depth for sorption (cm):	1
sediment oc (%):	5
sediment bulk density (kg/L):	0.8
ratio of field to water body:	10

Step 1: Input into surface water

Crop	Distance crop-water (m)	Drift (% of application)	Run-off/drainage (% of application)
cereals, spring	1	2.8	10
cereals, winter	1	2.8	10
citrus	3	15.7	10
cotton	1	2.8	10
field beans	1	2.8	10
grass / alfalfa	1	2.8	10
hops	3	19.3	10
legumes	1	2.8	10
maize	1	2.8	10
oil seed rape, spring	1	2.8	10
oil seed rape, winter	1	2.8	10
olives	3	15.7	10
pome / stone fruit, early applns	3	29.2	10
pome / stone fruit, late applns	3	15.7	10
potatoes	1	2.8	10
soybeans	1	2.8	10
sugar beet	1	2.8	10
sunflower	1	2.8	10
tobacco	1	2.8	10
vegetables, bulb	1	2.8	10
vegetables, fruiting	1	2.8	10
vegetables, leafy	1	2.8	10
vegetables, root	1	2.8	10
vines, early applns	3	2.7	10
vines, late applns	3	8.0	10
appln, aerial	3	33.2	10
appln, hand (crop < 50 cm)	1	2.8	10
appln, hand (crop > 50 cm)	3	8.0	10
no drift (incorp or seed trtmt)	1	0	10

Step 2: Input into surface water via spray drift

Crop	Distance crop-water	Number of application per season							
	(m)	1	2	3	4	5	6	7	>7
cereals, spring	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
cereals, winter	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
citrus	3	15.7	12.1	11.0	10.1	9.7	9.2	9.1	8.7
cotton	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
field beans	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
grass / alfalfa	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
hops	3	19.3	17.7	15.9	15.4	15.1	14.9	14.6	13.5
legumes	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
maize	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
oil seed rape, spring	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
oil seed rape, winter	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
olives	3	15.7	12.1	11.0	10.1	9.7	9.2	9.1	8.7
pome / stone fruit, (early)	3	29.2	25.5	24.0	23.6	23.1	22.8	22.7	22.2
pome / stone fruit (late)	3	15.7	12.1	11.0	10.1	9.7	9.2	9.1	8.7
potatoes	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
soybeans	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
sugar beet	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
sunflower	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
tobacco	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
vegetables, bulb	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
vegetables, fruiting	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
vegetables, leafy	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
vegetables, root	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
vines, early applns	3	2.7	2.5	2.5	2.5	2.4	2.3	2.3	2.3
vines, late applns	3	8.0	7.1	6.9	6.6	6.6	6.4	6.2	6.2
appln, aerial	3	33.2	33.2	33.2	33.2	33.2	33.2	33.2	33.2
appln, hand (crop < 50 cm)	1	2.8	2.4	2.0	1.9	1.8	1.6	1.6	1.5
appln, hand (crop > 50 cm)	3	8.0	7.1	6.9	6.6	6.6	6.4	6.2	6.2
no drift (incorp/seed trtmt)	1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Step 3: crop interception

Crop	no interception	minimal crop cover	intermediate crop cover	full canopy
BBCH-code*	00 - 09	10 - 19	20 - 39	40 - 89
cereals, spring	0	0.25	0.5	0.7
cereals, winter	0	0.25	0.5	0.7
citrus	0	0.7	0.7	0.7
cotton	0	0.3	0.6	0.75
Field beans	0	0.25	0.4	0.7
grass / alfalfa	0	0.4	0.6	0.75
Hops	0	0.2	0.5	0.7
legumes	0	0.25	0.5	0.7
maize	0	0.25	0.5	0.75
oil seed rape, spring	0	0.4	0.7	0.75
oil seed rape, winter	0	0.4	0.7	0.75
olives	0	0.7	0.7	0.7
pome / stone fruit, (early)	0	0.2	0.4	0.7
pome / stone fruit (late)	0	0.2	0.4	0.7
potatoes	0	0.15	0.5	0.7
soybeans	0	0.2	0.5	0.75
sugar beet	0	0.2	0.7	0.75
sunflower	0	0.2	0.5	0.75
tobacco	0	0.2	0.7	0.75
vegetables, bulb	0	0.1	0.25	0.4
vegetables, fruiting	0	0.25	0.5	0.7
vegetables, leafy	0	0.25	0.4	0.7
vegetables, root	0	0.25	0.5	0.7
Vines, early applns	0	0.4	0.5	0.7
Vines, late applns	0	0.4	0.5	0.7
appln, aerial	0	0.2	0.5	0.7
appln, hand (crop < 50 cm)	0	0.2	0.5	0.7
appln, hand (crop > 50 cm)	0	0.2	0.5	0.7
no drift (incorp/seed trtmt)	0	0	0	0

NOTE: indicative, adapted coding, the BBCH-codes mentioned do not exactly match (BBCH 1994)

Reference

BBCH 1994: *Compendium of growth stage indication keys for mono- and dicotyledonous plants - extended BBCH scale.* Ed R Stauss. Published by BBA, BSA, IGZ, IVA, AgrEvo, BASF, Bayer & Ciba, ISBN 3-9520749-0-X, Ciba-Geigy AG, Postfach, CH-4002 Basel, Switzerland.

APPENDIX III – DIFFICULT TO TEST SUBSTANCES

Adapted from the GHS (United Nations, 2003)

Valid aquatic toxicity tests require the dissolution of the test substance in the water media under the test conditions recommended by the guideline. In addition, a bioavailable exposure concentration should be maintained for the duration of the test. Some chemical substances are difficult to test in aquatic systems and guidance has been developed to assist in testing these materials.

The OECD provides a *Guidance Document on Aquatic Toxicity testing of Difficult Substances and Mixtures* (OECD, 2000). This document is a good source of information on the types of substances that are difficult to test and the steps needed to ensure valid conclusions from tests with these materials and assessors should be familiar with this document.

Nevertheless, much test data exist that may have used testing methodologies which, while not in conformity with what might be considered best practice today, can still yield suitable information. Such data require special guidance on interpretation, although expert judgement must ultimately be used in determining data validity. Such difficult to test substances may be poorly soluble, volatile, or subject to rapid degradation due to such processes as photo transformation, hydrolysis, oxidation, or biotic degradation. When testing algae, coloured materials may interfere with the test end-point by attenuating the light needed for cell growth. In a similar manner, substances tested as cloudy dispersions above solubility may give rise to false toxicity measurements. Loading of the water column with test material can be an issue for particulates or solids such as metals. Petroleum distillate fractions can also pose loading problems, as well as difficult interpretational problems when deciding on the appropriate concentrations for determining L(E)C50 values. The *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures* describes the more common properties of many types of substances that are likely to pose testing difficulties.

Stability: If test chemical concentrations are expected to fall below 80% of nominal, then testing - in order to be valid - may require exposure regimes that provide for renewal of the test material. In this case, semi-static or flow-through conditions are preferred. Therefore, problems arise when testing on algae where the standard guidelines generally include static tests to be conducted. While alternative exposure regimes are possible for crustacea and fish, these tests are frequently conducted on static conditions as included in the internationally agreed guidelines. In these tests, a certain level of degradation as well as other relevant factors has to be tolerated and appropriate account must be taken in calculations of toxic concentrations. Where degradation occurs, it is also important to consider the influence of the toxicity of the degradation products on the recorded toxicity in the test. Expert judgement will need to be exercised when deciding if these data can be used.

Degradation: When a compound breaks down or degrades under test conditions, then expert judgement should be used in calculating toxicity, including consideration of known or likely breakdown products. Concentrations of the parent material and all significant toxic degradates are desirable parameters. If degradates are expected to be relatively non-toxic, then renewable exposure regimes are desirable in order to ensure that levels of the parent compounds are maintained.

Saturation: For single component substances, interpretation should be based only on toxic responses observed in the soluble range, and not on total chemical loading above solubility. Frequently, data are available which indicate toxicity at levels in excess of water solubility and, while these data will often be regarded as not valid, some interpretation may be possible. These problems generally apply when testing poorly soluble substances, and guidance on how to interpret such data is provided in the *Guidance Document on Aquatic Toxicity testing of Difficult Substances and Mixtures*.

Perturbation of test media: Special provisions may be needed to ensure dissolution of difficult to test substances. Such measures should not lead to significant changes in the test media when such changes are likely to lead to an increase or decrease in the apparent toxicity and hence the classification level of the test substance.

Complex substances: Measurement of exposure concentrations of mixtures is often difficult, and in some cases impossible. Substances such as petroleum distillate fractions, polymers, substances with significant levels of impurities, etc can pose special problems since the toxic concentration is difficult to define and impossible to verify. Typical testing procedures often rely on the formation of a water soluble fraction (WSF) or water accommodated fraction (WAF) and data are reported in terms of loading rates.

It is desirable to have stabilised and analytically measured test concentrations. Although measured concentrations are preferred, interpretation may be based on nominal concentration studies when these are the only valid data available under certain circumstances. If the material is likely to substantially degrade or otherwise be lost from the water column, care must be taken in data interpretation in order to take the loss of the toxicant during the test into account, if relevant and possible.

The following paragraphs provide some detailed guidance on some of these interpretational problems. In doing so it should be remembered that this is guidance and therefore hard and fast rules cannot be applied. The nature of many of the difficulties mean that expert judgement must always be applied both in determining whether there is sufficient information in a test for a judgement to be made on its validity, and also whether a toxicity level can be determined that is suitable for use in applying the classification criteria.

Unstable substances

While testing procedures that minimise the impacts of instability in the test media should ideally have been adopted, in practice it can be almost impossible to maintain a concentration throughout the test in certain cases. Common causes of such instability are oxidation, hydrolysis, photodegradation and biodegradation. While the latter forms of degradation can be more readily controlled, such controls are frequently absent in much existing testing. Nevertheless, for some testing, particularly acute and chronic fish toxicity testing, a choice of exposure regimes is available to help minimise losses due to instability. This should be taken into account in deciding on the test data validity.

Where instability is a factor in determining the level of exposure during the test, an essential prerequisite for data interpretation is the existence of measured exposure concentrations at suitable time points throughout the test. In the absence of analytically measured concentrations at least at the start and end of test, no valid interpretation can be made and the test should be considered as invalid for classification purposes and may be invalid for risk assessment purposes. Where measured data are available, a number of practical rules can be considered by way of guidance in interpretation:

- Where measured data are available for the start and end of test (as is normal for the acute *Daphnia* and algal tests), the L(E)C50 may be calculated based on the geometric mean of the start and end of test concentrations. Where the concentrations at the end of the test are below the analytical detection limit, such concentrations shall be considered to be half that detection limit.
- Where measured data are available at the start and end of media renewal periods (as may be available for the semi-static tests), the geometric mean for each renewal period should be calculated, and the mean exposure over the whole exposure period calculated from these data.
- Where the toxicity can be attributed to a degradation breakdown product, and the concentrations of this product are known, the L(E)C50 may be calculated based on the geometric mean of the degradation product concentration, back calculated to the parent substance.
- Similar principles may be applied to measured data in chronic toxicity testing.

Poorly soluble substances

These substances, usually taken to be those with solubility in water of < 1 mg/L, are frequently difficult to dissolve in the test media. As well, the dissolved concentrations will often prove difficult to measure at the low concentrations anticipated. For many substances, the true solubility in the test media will be unknown, and will often be recorded as < detection limit in purified water. Nevertheless such substances can show toxicity. Where no toxicity is found, judgement must be applied to determine whether the result can be considered valid. Judgement should err on the side of caution and should not underestimate the hazard.

Ideally, tests using appropriate dissolution techniques and with accurately measured concentrations within the range of water solubility should be used. Where such test data are available, they should be used in preference to other data. It is normal, however, particularly when considering older data, to find such substances with toxicity levels recorded in excess of the water solubility, or where the dissolved levels are below the detection limit of the analytical method. Thus, in both circumstances, it is not possible to verify the actual exposure concentrations using measured data. Where these are the only data available, some practical rules can be considered by way of general guidance:

- Where the acute toxicity is recorded at levels in excess of the water solubility, the L(E)C50 may be considered to be equal to or below the measured water solubility. In making this decision, due attention should be paid to the possibility that the excess undissolved substance may have given rise to physical effects on the test organisms. Where this is considered the likely cause of the effects observed, the test should be considered invalid.

- Where no acute toxicity is recorded at levels in excess of the water solubility, the L(E)C50 may be considered to be greater than the measured water solubility. In making a decision that the substance shows no acute toxicity, due account should be taken of the techniques used to achieve the maximum dissolved concentrations. Where these are not considered as adequate, the test should be considered invalid.
- Where the water solubility is below the detection limit of the analytical method for a substance, and acute toxicity is recorded, the L(E)C50 may be considered to be less than the analytical detection limit. Where no toxicity is observed, the L(E)C50 may be considered to be greater than the water solubility. Due consideration should also be given to the quality criteria mentioned above.
- Where chronic toxicity data are available, the same general rules should apply. In principle, only data showing no effects at the water solubility limit, or greater than 1 mg/L need be considered. Again, where these data cannot be validated by consideration of measured concentrations, the techniques used to achieve the maximum dissolved concentrations must be considered as appropriate.

Other factors contributing to concentration loss

A number of other factors can also contribute to losses in concentration. While some can be avoided by correct study design, interpretation of data where these factors have contributed may be necessary.

- **Sedimentation**: this can occur during a test for a number of reasons. A common explanation is that the substance has not truly dissolved despite the apparent absence of particulates, and agglomeration occurs during the test leading to precipitation. In these circumstances, the L(E)C50 may be considered to be based on the end of test concentrations. Equally, precipitation can occur through reaction with the media.
- **Adsorption**: this can occur for substances of high adsorption characteristics such as high log Kow substances. Where this occurs, the loss of concentration is usually rapid and exposure may best be characterised by the end of test concentrations.
- **Bioaccumulation**: losses may occur through the bioaccumulation of a substance into the test organisms. This may be particularly important where the water solubility is low and log Kow correspondingly high. The L(E)C50 may be calculated based on the geometric mean of the start and end of test concentrations.

Perturbation of the test media

Polymers are typically not available in aquatic systems. Dispersible polymers and other high molecular mass materials can perturb the test system and interfere with uptake of oxygen. This can give rise to mechanical or secondary effects. These factors need to be taken into account when considering data from these substances. Many polymers behave like complex substances through having a significant low molecular mass fraction which can leach from the bulk polymer. This is considered further below.

Complex substances

Complex substances are characterised by a range of chemical structures, frequently in a homologous series, but covering a wide range of water solubilities and other physico-chemical characteristics. Upon addition to water, equilibrium will be reached between the dissolved and undissolved fractions that will be characteristic of the loading of the substance. For this reason, such complex substances are usually tested as a WSF or WAF, and the L(E)C50 recorded based on the loading or nominal concentrations. Analytical support data are not normally available since the dissolved fraction will itself be a complex mixture of components. The toxicity parameter is sometimes referred to as LL50, related to the lethal loading level. This loading level from the WSF or WAF may be used directly in the classification criteria.

Polymers represent a special kind of complex substance, requiring consideration of the polymer type and their dissolution/dispersal behaviour. Polymers may dissolve as such without change (true solubility related to particle size), be dispersible, or portions consisting of low molecular weight fractions may go into solution. In the latter case, the testing of a polymer is a test of the ability of low molecular mass material to leach from the bulk polymer, and whether this leachate is toxic. Thus, it can be considered in the same way as a complex mixture in that a loading of polymer can best characterise the resultant leachate and, hence, the toxicity can be related to this loading.

APPENDIX IV – ENDOCRINE DISRUPTION POTENTIAL

Endocrine disruption caused by a chemical should be viewed as an effect that can be assessed within the normal framework outlined for effects assessment. However, the main difficulty with this effect is the lack of test data available to adequately support assessment. Some information may be gleaned from studies received as part of the data package, for example, long-term studies on fish, aquatic invertebrates, birds and some mammal studies. However, the extent to which these studies provide useful information is questionable (see below).

To address this problem, the OECD has established the Endocrine Disruptors Testing and Assessment Taskforce. Documents related to this activity may be found at <http://www.oecd.org/document/62/0,2340,en_2649_34377_2348606_1_1_1_1,00.html> and assessors should be familiar with developments in this area. One useful report available from this site is the *Detailed Review Paper: Appraisal of Test Methods For Sex Hormone Disrupting Chemicals* (OECD, 2002). The content of this document is summarised below:

In recent years several national and international workshops have concluded that chemicals present in the environment may be exerting an adverse effect on human and wildlife reproductive health, and a number of environmental contaminants have shown oestrogenic or related activities in laboratory studies. Although there is as yet no evidence to suggest a causal link between these observations, many gaps in our knowledge have been identified. It has been recognised that there is an urgent need to establish validated *in vivo* and *in vitro* screening assays to test for the oestrogenic and androgenic activities of chemicals.

OECD 2002, details and critically assesses the ability of existing, relevant OECD test methods to detect a chemical's sex hormone-disrupting potential with regard to reproductive processes. Assessors should be familiar with the content of this document when using current standard tests to evaluate the potential for endocrine disruption impacts. This reference also reviews a range of non-regulatory model systems that have been used in scientific research in order to assess their suitability. Consideration of other forms of endocrine disruption (involving other hormonal systems) and of non-reproductive functions of the sex hormones is outside the scope of this document.

Potential modifications of, or additions to, existing regulatory test methods are identified, as are other models considered suitable for formal validation and adoption within an expanded testing battery. Possible approaches to the routine assessment of chemicals for sex hormone-disrupting potential, and gaps in scientific knowledge are also discussed.

Limitations of current OECD test methods

None of the existing OECD test methods were specifically designed to detect the endocrine disruptive activity of chemicals. Indeed, methods for assessing acute toxicity (eg in fish, OECD 202; earthworm, OECD 207; or rodent, OECD 401) and those designed to investigate a specific type of toxicity (eg neurotoxicity, OECD 424) are unlikely to be suited to detecting such processes. Nonetheless, some of the vertebrate designs might be expected to be able to detect whether a chemical has significant sex hormone-disrupting activity, either in their current state or with a degree of modification. However, additional studies might then be required to clarify the mechanism of action. In the case of the mammalian studies, several gaps in design can be identified, including the degree of pathological examination of the gonads and secondary sex organs and, for reproductive studies, detailed examination of offspring.

There are very few existing OECD tests in non-vertebrates. Only the reproductive study in daphnids might be expected to be of help in detecting sex hormone-disruption. However, given the different endocrine systems and the wide range of reproductive strategies found among the invertebrate taxa, this test alone would provide insufficient information for hazard assessment. Consequently, there is an urgent need to develop new test models for invertebrates.

Possible modifications of existing OECD test methods

For testing of vertebrates - at least in the short term - the most promising approach seems to be to enhance the existing OECD test guidelines. A number of possible enhancements have been identified, many of which are most relevant to the mammalian test designs. These include:

- extension of organ weight and histopathology requirements for gonads and accessory sex organs
- pathological examination of offspring, where appropriate
- measurement of sex hormone blood levels
- detailed assessment of spermatogenesis and/or semen quality

- monitoring of oestrus cyclicity
- enhancement of current monitoring of physical and behavioural development, and of learning and memory functions in offspring
- possibly, investigation of accessory sex organ secretory products.

As noted above, for the study of invertebrate species only the existing test in daphnids might be able to be enhanced to provide more information on sex hormone-disruptive activity. The sensitivities of the possible end-points need to be established, as well as their importance as markers of toxic hazard. Pragmatically, emphasis might first be given to enhancing the sub-chronic test designs so as to maximise benefit in terms of numbers of chemicals screened. Nonetheless, unless the end-points used in such designs can be shown to be predictive of adverse effects at all key life stages (eg during early development), it will be essential to optimise the sensitivity of existing reproductive study designs and, where appropriate, to develop new models in a range of taxa.

Non-regulatory test models proposed for further development and adoption

In addition to enhancing existing test designs, it is appropriate to consider the development of other toxicity and reproductive tests in a range of species so as to be able to better assess ecological hazard. This is likely to involve development of a range of multi-generation designs in a range of species - including fish and invertebrates - that could be used for regulatory assessment. Although designs developed by various regulatory agencies throughout the world are likely to be worthy of consideration, it is felt that there will be a need to conduct basic research before an adequate test battery can be developed.

Given the number of chemicals that potentially require testing, there is also a need for a range of simple *in vivo* and *in vitro* models that, even if unable to provide sufficient confidence on which to base regulatory decisions, would nonetheless enable initial screening and prioritisation of chemicals or would be of value in elucidating mechanisms. The following non-regulatory models are proposed for further development, with a view to possible adoption in screening designs:

- the rat vaginotrophic and/or uterotrophic assays to assess potential interference with the oestrogenic hormonal system (there is a need to optimise study designs and perform cross comparisons to determine which are the most suitable for progression)
- the prostate weight of castrated rats as a marker for androgenic hormone modulation
- assay of vitellogenin in males of oviparous species, which might be a useful biomarker of exposure to oestrogens whilst, as an interim measure, the use of models involving changes in secondary sexual morphology of fish might be appropriate.

Work is also required on developing suitable non-vertebrate models that can detect disruption of endocrine systems having no mammalian correlate, such as the arthropod hormone - ecdysone and juvenile hormone.

At present, it is not possible to recommend the formal adoption of any of the *in vitro* assays because of the various limitations and difficulties inherent in the current designs. These include: *in vitro* end-points that are dependent on specific receptor or response element interactions, which may not mimic *in vivo* modes of action; the inability of many systems to distinguish agonists from antagonists; and the finding that existing *in vitro* models lack satisfactory metabolic systems or may show only limited chemical uptake. Significant interlaboratory differences in specificity, sensitivity and reproductibility also exist for some systems, and the significance of *in vitro* findings must be translated to intact organisms where physiological processes may play critical roles in determining activity. There is also a need to establish the predictability and sensitivity of such models against an appropriate gold standard *in vivo* methodology. Further development is recommended of assays using:

- human cell lines, such as Ishikawa and MCF-7 cell lines for oestrogenic activity
- yeast cells for oestrogenicity and androgenicity
- the trout hepatocyte vitellogenin assay for oestrogenic activity.

Continued development of structure-activity relationship models is also recommended.

Requirements for basic research

A number of general research issues of importance to the understanding of endocrine-disruptive activity have been identified from this review. These include: clarification of basic mechanisms of action of endocrine disrupters (especially the importance of non-nuclear and nuclear receptor-mediated effects and of interactions which are not receptor-mediated; and cross-species differences in pharmacokinetics and pharmacodynamics and clarification of the predictability of effects *in utero/ovo* from data generated using adult forms. Resolution of such questions would be assisted by agreement on a set of reference chemicals and the continued study of structure-

activity relationships. There is also a need to rank the relative sensitivity and predictivity of the various end-points that could be used to assess endocrine disruption, in order to facilitate selection of those end-points most appropriate for inclusion in regulatory test designs. Basic knowledge of comparative endocrinology needs to be enhanced to facilitate the identification of suitable species (including representative invertebrates) for inclusion in an expanded range of wildlife tests: this is necessary in order to permit adequate screening for end-points and processes having no mammalian correlate.

With specific reference to the issue of sex hormone-disruption, the following principal research recommendations are made:

- rank sensitivities of the various marker end-points, and assess their relevance and relative importance as markers of endocrine-mediated toxicity
- establish a reference set of chemicals of defined activity to assist in method development and validation
- elucidate the dose-response profiles for endocrine-disruptive mechanisms and apply them to dosage selection during testing
- assess whether classical toxicological assumptions of cross-group predictivity hold true for endocrine-disruptive mechanisms
- develop simple, inexpensive models spanning a range of ecologically-relevant species, focusing on processes with no mammalian correlate
- assess the extent to which interactions occur between sex hormone-disrupting chemicals in mixtures, particularly in vivo, and investigate the likely frequency of occurrence of such interactions between different sex hormone-disrupters.

Other more general research needs have also been identified and are discussed in detail in OECD, 2002.

APPENDIX V – BIOACCUMULATION

Adapted from the GHS (United Nations, 2003). For further information, the source document should be consulted.

Introduction

Bioaccumulation is one of the important intrinsic properties of chemical substances that determine the potential environmental hazard. Bioaccumulation of a substance into an organism is not a hazard in itself, but bioconcentration and bioaccumulation will result in a body burden, which may potentially lead to toxic effects. Assessors should be aware of the distinction between bioconcentration and bioaccumulation. Here bioconcentration is defined as the net result of uptake, transformation, and elimination of a substance in an organism due to waterborne exposure, whereas bioaccumulation includes all routes of exposure (ie via air, water, sediment/soil, and food). Finally, biomagnification is defined as accumulation and transfer of substances via the food chain, resulting in an increase of internal concentrations in organisms on higher levels of the trophic chain. For most organic chemicals uptake from water (bioconcentration) is believed to be the predominant route of uptake. Only for very hydrophobic substances does uptake from food becomes important. Also, the harmonised classification criteria use the bioconcentration factor (or the octanol/water partition coefficient) as the measure of the potential for bioaccumulation. For these reasons, the present guidance document only considers bioconcentration and does not discuss uptake via food or other routes.

Apart from the chemical's intrinsic properties, the degree of bioconcentration also depends on factors such as the degree of bioavailability, the physiology of test organism, maintenance of constant exposure concentration, exposure duration, metabolism inside the body of the target organism and excretion from the body. The interpretation of the bioconcentration potential therefore requires an evaluation of the intrinsic properties of the substance, as well as of the experimental conditions under which bioconcentration factor (BCF) has been determined. A decision scheme for application of bioconcentration or log Kow data for classification purposes has been developed. The emphasis of this appendix is organic substances and organo-metals.

Data on bioconcentration properties of a substance may be available from standardised tests or may be estimated from the structure of the molecule. The interpretation of such bioconcentration data often requires detailed evaluation of test data. To this end, assessors should be familiar with Appendices III and IV of Annex 8 of the GHS as these outline the basic principles of the experimental and estimation methods for determination of BCF and Kow of organic substances and give guidance on the influence of external and internal factors on the bioaccumulation potential or organic substances.

Interpretation of bioconcentration data

Bioconcentration of an organic substance can be experimentally determined in bioconcentration experiments, during which BCF is measured as the concentration in the organism relative to the concentration in water under steady-state conditions and/or estimated from the uptake rate constant and the elimination rate constant. In general, the potential of an organic substance to bioconcentrate is primarily related to the lipophilicity of the substance. A measure of lipophilicity is the n-octanol/water partition coefficient (Kow) that, for lipophilic non-ionic organic substances, undergoing minimal metabolism or biotransformation within the organism, is correlated with the bioconcentration factor. Therefore, Kow is often used for estimating the bioconcentration of organic substances, based on the empirical relationship between log BCF and log Kow. For most organic substances, estimation methods are available for calculating the Kow. Data on the bioconcentration properties of a substance may thus be (i) experimentally determined, (ii) estimated from experimentally determined Kow, or (iii) estimated from Kow values derived by use of Quantitative Structure Activity Relationships (QSARs). Guidance for interpretation of such data is given below together with guidance on assessment of chemical classes, which need special attention.

Bioconcentration factor (BCF*)*

The bioconcentration factor is defined as the ratio on a weight basis between the concentration of the chemical in biota and the concentration in the surrounding medium, here water, at steady state. BCF can thus be experimentally derived under steady-state conditions, on the basis of measured concentrations. However, BCF can also be calculated as the ratio between the first-order uptake and elimination rate constants; a method which does not require equilibrium conditions.

Different test guidelines for the experimental determination of bioconcentration in fish have been documented and adopted, the most generally applied being the OECD test guideline (OECD 305, 1996).

Experimentally derived BCF values of high quality are ultimately preferred as such data override surrogate data, eg. Kow.

High quality data are defined as data where the validity criteria for the test method applied are fulfilled and described, eg maintenance of constant exposure concentration; oxygen and temperature variations, and documentation that steady-state conditions have been reached, etc. The experiment will be regarded as a high-quality study, if a proper description is provided (eg by good laboratory practice) allowing verification that validity criteria are fulfilled. In addition, an appropriate analytical method must be used to quantify the chemical and its toxic metabolites in the water and fish tissue.

BCF values of low or uncertain quality may give a false and too low BCF value, for example, application of measured concentrations of the test substance in fish and water, but measured after a too short an exposure period in which steady-state conditions have not been reached (cf. OECD 306, 1996, regarding estimation of time to equilibrium). Therefore, such data should be carefully evaluated before use and consideration should be given to using Kow instead.

If there is no BCF value for fish species, high-quality data on the BCF value for other species may be used (eg BCF determined on blue mussel, oyster, scallop (ASTM E 1022-94). Reported BCFs for microalgae should be used with caution.

For highly lipophilic substances, eg with log Kow above 6, experimentally derived BCF values tend to decrease with increasing log Kow. Conceptual explanations of this non-linearity mainly refer to either reduced membrane permeation kinetics or reduced biotic lipid solubility for large molecules. A low bioavailability and uptake of these substances in the organism will thus occur. Other factors comprise experimental artefacts, such as equilibrium not being reached, reduced bioavailability due to sorption to organic matter in the aqueous phase, and analytical errors. Special care should thus be taken when evaluating experimental data on BCF for highly lipophilic substances as these data will have a much higher level of uncertainty than BCF values determined for less lipophilic substances.

BCF in different test species

BCF values used for classification are based on whole body measurements. The optimal data for classification are BCF values derived using the OECD 305 test method or internationally equivalent methods, which uses small fish. Due to the higher gill surface to weight ratio for smaller organisms than larger organisms, steady-state conditions will be reached sooner in smaller organisms than in larger ones. The size of the organisms (fish) used in bioconcentration studies is thus of considerable importance in relation to the time used in the uptake phase, when the reported BCF value is based solely on measured concentrations in fish and water at steady-state. Thus, if large fish, for example, adult salmon, have been used in bioconcentration studies, it should be evaluated whether the uptake period was sufficiently long for steady state to be reached or to allow for a kinetic uptake rate constant to be determined precisely.

Furthermore, when using existing data, it is possible that the BCF values could be derived from several different fish or other aquatic species (eg clams) and for different organs in the fish. Thus, to compare these data to each other and to the criteria, some common basis or normalisation will be required. It has been noted that there is a close relationship between the lipid content of a fish or an aquatic organism and the observed BCF value. Therefore, when comparing BCF values across different fish species or when converting BCF values for specific organs to whole body BCFs, the common approach is to express the BCF values on common lipid content. If eg whole body BCF values or BCF values for specific organs are found in the literature, the first step is to calculate the BCF on a % lipid basis using the relative content of fat in the fish (cf. literature/test guideline for typical fat content of the test species) or the organ. In the second step the BCF for the whole body for a typical aquatic organism (ie small fish) is calculated assuming a common default lipid content. A default value of 5% is most commonly used as this represents the average lipid content of the small fish used in OECD 305 (1996).

Use of radiolabelled substances

The use of radiolabelled test substances can facilitate the analysis of water and fish samples. However, unless combined with a specific analytical method, the total radioactivity measurements potentially reflect the presence of the parent substance as well as possible metabolite(s) and possible metabolised carbon, which have been incorporated in the fish tissue in organic molecules. BCF values determined by use of radiolabelled test substances are, therefore, normally overestimated.

When using radiolabelled substances, the labeling is most often placed in the stable part of the molecule, for which reason the measured BCF value includes the BCF of the metabolites. For some substances it is the metabolite which is the most toxic and which has the highest bioconcentration potential. Measurements of the parent substance as well as the metabolites may thus be important for the interpretation of the aquatic hazard (including the bioconcentration potential) of such substances.

In experiments where radiolabelled substances have been used, high radiolabel concentrations are often found in the gall bladder of fish. This is interpreted as caused by biotransformation in the liver and subsequently by excretion of metabolites in the gall bladder. When fish do not eat, the content of the gall bladder is not emptied into the gut, and high concentrations of metabolites may build up in the gall bladder. The feeding regime may thus have a pronounced effect on the measured BCF. In the literature many studies are found where radiolabelled compounds are used, and where the fish are not fed. As a result high concentrations of radioactive material are found in the gall bladder. In these studies the bioconcentration may in most cases have been overestimated. Thus, when evaluating experiments, in which radiolabelled compounds are used, it is essential to evaluate the feeding regime as well.

If the BCF in terms of radiolabelled residues is documented to be ≥1000, identification and quantification of degradation products, representing ≥10% of total residues in fish tissues at steady-state, are for example, pesticides, strongly recommended in the OECD guideline No. 305 (1996). If no identification and quantification of metabolites are available, the assessment of bioconcentration should be based on the measured radiolabelled BCF value. If, for highly bioaccumulative substances (BCF ≥500), only BCFs based on the parent compound and on radiolabelled measurements are available, the latter should thus be used in relation to classification.

Octanol-water-partitioning coefficient (K_{ow})

For organic substances experimentally derived high-quality Kow values, or values which are evaluated in reviews and assigned as the 'recommended values', are preferred over other determinations of Kow. When no experimental data of high quality are available, validated quantitative structure activity relationships (QSARs) for log Kow may be used in the classification process. Such validated QSARs may be used without modification to the agreed criteria if they are restricted to chemicals for which their applicability is well characterised. For substances like strong acids and bases, substances which react with the eluent, or surface-active substances, a QSAR estimated value of K_{ow} or an estimate based on individual n-octanol and water solubilities should be provided instead of an analytical determination of Kow. Measurements should be taken on ionisable substances in their non-ionised form (free acid or free base) only by using an appropriate buffer with pH below pK for free acid or above the pK for free base.

Chemical classes that need special attention with respect to BCF and K_{ow} values

There are certain physico-chemical properties, which can make the determination of BCF or its measurement difficult. These may be substances, which do not bioconcentrate in a manner consistent with their other physico-chemical properties, for example, steric hindrance or substances which make the use of descriptors inappropriate, eg surface activity, which makes both the measurement and use of log Kow inappropriate.

Difficult substances

Some chemical substances are difficult to test in aquatic systems and guidance has been developed to assist in testing these materials (see Appendix III above).

Difficult to test substances may be poorly soluble, volatile, or subject to rapid degradation due to such processes as phototransformation, hydrolysis, oxidation, or biotic degradation.

To bioconcentrate organic compounds, a substance needs to be soluble in lipids, present in the water, and available for transfer across the fish gills. Properties that alter this availability will thus change the actual bioconcentration of a substance, when compared with the prediction. For example, readily biodegradable substances may only be present in the aquatic compartment for short periods of time. Similarly, volatility, and hydrolysis will reduce the concentration and the time during which a substance is available for bioconcentration. A further important parameter, which may reduce the actual exposure concentration of a substance, is adsorption, either to particulate matter or to surfaces in general. There are a number of substances, which have shown to be rapidly transformed in the organism, thus, leading to a lower BCF value than expected. Substances that form micelles or aggregates may bioconcentrate to a lower extent than would be predicted from simple physico-chemical properties. This is also the case for hydrophobic substances that are contained in micelles formed as a consequence of the use of dispersants. Therefore, the use of dispersants in bioaccumulation tests is discouraged.

In general, for difficult to test substances, measured BCF and Kow values – based on the parent substance – are a prerequisite for the determination of the bioconcentration potential. Furthermore, proper documentation of the test concentration is a prerequisite for the validation of the given BCF value.

Poorly soluble and complex substances

Special attention should be paid to poorly soluble substances. Frequently the solubility of these substances is recorded as less than the detection limit, which creates problems in interpreting the bioconcentration potential.

For such substances the bioconcentration potential should be based on experimental determination of log Kow or QSAR estimations of log Kow. When a multi-component substance is not fully soluble in water, it is important to attempt to identify the components of the mixture as far as practically possible and to examine the possibility of determining its bioaccumulation potential using available information on its components. When bioaccumulating components constitute a significant part of the complex substance (eg more than 20% or for hazardous components an even lower content), the complex substance should be regarded as being bioaccumulating.

High molecular weight substances

Above certain molecular dimensions, the potential of a substance to bioconcentrate decreases. This is possibly due to steric hindrance of the passage of the substance through gill membranes. It has previously been proposed that a cut-off limit of 700 for the molecular weight could be applied. However, this cut-off has been subject to criticism and an alternative cut-off of 1000 has been proposed in relation to exclusion of consideration of substances with possible indirect aquatic effects. In general, bioconcentration of possible metabolites or environmental degradation products of large molecules should be considered. Data on bioconcentration of molecules with a high molecular weight should therefore be carefully evaluated and only be used if such data are considered to be fully valid in respect to both the parent compound and its possible metabolites and environmental degradation products.

Surface-active agents

Surfactants consist of a lipophilic (most often an alkyl chain) and a hydrophilic part (the polar headgroup). According to the charge of the headgroup, surfactants are subdivided into classes of anionic, cationic, non-ionic, or amphoteric surfactants. Due to the variety of different headgroups, surfactants are a structurally diverse class of compounds, which is defined by surface activity rather than by chemical structure. The bioaccumulation potential of surfactants should thus be considered in relation to the different subclasses (anionic, cationic, non-ionic, or amphoteric) instead of to the group as a whole. Surface-active substances may form emulsions, in which the bioavailability is difficult to ascertain. Micelle formation can result in a change of the bioavailable fraction even when the solutions are apparently formed, thus, giving problems in interpretation of the bioaccumulation potential**.**

Experimentally derived bioconcentration factors

Measured BCF values on surfactants show that BCF may increase with increasing alkyl chain length and be dependant of the site of attachment of the head group, and other structural features.

Octanol-water-partition coefficient (K_{ow})

The octanol-water partition coefficient for surfactants cannot be determined using the shakeflask or slow stirring method because of the formation of emulsions. In addition, the surfactant molecules will exist in the water phase almost exclusively as ions, whereas they will have to pair with a counter-ion in order to be dissolved in octanol. Therefore, experimental determination of Kow does not characterise the partition of ionic surfactants. On the other hand, it has been shown that the bioconcentration of anionic and non-ionic surfactants increases with increasing lipophilicity. It has further been shown that for some surfactants, an estimated log Kow value could represent the bioaccumulation potential; however, for other surfactants some 'correction' to the estimated log Kow value using the method was required. These results illustrate that the quality of the relationship between log Kow estimates and bioconcentration depends on the class and specific type of surfactants involved. Therefore, the classification of the bioconcentration potential based on log Kow values should be used with caution.

Conflicting data and lack of data

Conflicting BCF data

In situations where multiple BCF data are available for the same substance, the possibility of conflicting results might arise. In general, conflicting results for a substance, which has been tested several times with an appropriate bioconcentration test, should be interpreted by a weight of evidence approach. This implies that if experimental determined BCF data, both ≥ and <500, have been obtained for a substance the data of the highest quality and with the best documentation should be used for determining the bioconcentration potential of the substance. If differences still remain, if eg high-quality BCF values for different fish species are available, generally the highest valid value should be used as the basis for assessment and classification. When larger data sets (four or more values) are available for the same species and life stage, the geometric mean of the BCF values may be used as the representative BCF value for that species.

Conflicting log K_{ow} data

The situations, where multiple log Kow data are available for the same substance, the possibility of conflicting results might arise. For example, if log Kow data both ≥ and <4 have been obtained for a substance, then the data of the highest quality and the best documentation should be used for determining the bioconcentration potential of the substance. If differences still exist, generally the highest valid value should take precedence. In such situation, QSAR estimated log Kow could be used as a guide.

Expert judgement

If no experimental BCF or log Kow data or no predicted log Kow data are available, the potential for bioconcentration in the aquatic environment may be assessed by expert judgement. This may be based on a comparison of the structure of the molecule with the structure of other substances for which experimental bioconcentration or log Kow data or predicted Kow are available.

Decision scheme

Based on the above discussions and conclusions, a decision scheme has been elaborated which may facilitate decisions as to whether or not a substance has the potential for bioconcentration in aquatic species.

Experimentally derived BCF values of high quality are ultimately preferred for classification purposes. BCF values of low or uncertain quality should not be used for classification purposes if data on log Kow are available because they may give a false and too low BCF value, for example, due to a too short exposure period in which steady-state conditions have not been reached. If no BCF is available for fish species, high quality data on the BCF for other species (eg mussels) may be used.

For organic substances, experimentally derived high quality Kow values, or values which are evaluated in reviews and assigned as the recommended values, are preferred. If no experimentally data of high quality are available validated Quantitative Structure Activity Relationships (QSARs) for log Kow may be used in the classification process. Such validated QSARs may be used without modification, if restricted to chemicals for which their applicability is well characterised. For substances like strong acids and bases, metal complexes, and surface-active substances a QSAR estimated value of Kow or an estimate based on individual *n*-octanol and water solubilities should be provided instead of an analytical determination of Kow.

If data are available but not validated, expert judgement should be used.

Whether or not a substance has a potential for bioconcentration in aquatic organisms could thus be decided in accordance with the following scheme:

Valid/high quality experimentally determined BCF value → YES:

→ BCF ≥500: *The substance has a potential for bioconcentration*

→ BCF <500: *The substance does not have a potential for bioconcentration.*

Valid/high quality experimentally determined BCF value → NO:

→ Valid/high quality experimentally determined log K_{ow} value → YES:

→ log K_{ow} ≥4*: The substance has a potential for bioconcentration*

→ log K_{ow} <4: *The substance does not have a potential for bioconcentration.*

Valid/high quality experimentally determined BCF value → NO:

→ Valid/high quality experimentally determined log K_{ow} value → NO:

→ Use of validated QSAR for estimating a log K_{ow} value → YES:

→ log K_{ow} ≥4: *The substance has a potential for bioconcentration*

→ log K_{ow} <4: The substance does not have a potential for bioconcentration.

Air-breathing animals

The above discussion, as paraphrased from the GHS, focuses on bioaccumulation in the aquatic compartment. There is an emerging area of bioaccumulation assessment that focuses on air-breathing animals. Modelling by Kelly *et al* (2007) shows that air breathing organisms exhibit higher biomagnification factors (BMFs) than water respiring organisms because of their greater ability to absorb and digest their diet, which is related to differences

in digestive tract physiology and body temperature. The model also shows that the relationship between the BMF and chemical properties is controlled by the rate of elimination. In water respiring organisms, elimination becomes sufficiently slow to cause biomagnification if the Kow of the chemical exceeds 105. For the air breathing organisms they studied, this occurs for chemicals with a high Koa (octanol-air partition co-efficient) of >106, which causes slow respiratory elimination, and a Kow of >102, causing slow elimination in urine or nitrogenous waste. Kelly *et al* (2007) postulate that the differences in biomagnification behaviour between air-breathing and water-respiring organisms implies that, for substances with a Koa of >106 and a Kow of >102, Kow and the BCF in fish are not good predictors of biomagnification in air-breathing animals.

The authors further found that application of their bioaccumulation model to identify potentially bioaccumulative substances among commercial chemicals reveals distinct differences in the biomagnification behaviour of chemicals in different food webs as follows:

Piscivorous food web:

- Concentrations of non-metabolising chemicals with Kow between 105 and 108 biomagnify in top level predatory fish up to 100-fold.
- No biomagnification occurs for less hydrophobic chemicals with Kow <105, which are efficiently eliminated by respiration, or for superhydrophobic organic substances with Kow >108, which are absorbed at very slow rates.

Marine mammalian food web - includes water-respiring invertebrates and fish and air-breathing birds and mammals:

- Poorly metabolising chemicals with a Kow >105 and Koa >106 biomagnify, attaining concentrations in top predators (polar bears) up to 10,000 times the concentrations in primary producers.
- Less hydrophobic chemicals with Kow <105 and Koa >106 also biomagnify strongly, with concentrations in polar bears exceeding those in primary producers by up to 3000-fold.
- Chemicals with Kow <102 do not biomagnify in this food web regardless of their high Koa because air-breathing animals eliminate them through urinary excretion.

Terrestrial food webs:

- Chemicals with a Kow between 102 and 1010 and a Koa >106 can biomagnify up to 400-fold if not metabolised
- Chemicals with a Kow between 103 and 109 achieve a similar degree of biomagnification, given the same Koa.

APPENDIX VI – HIGHER TIER AQUATIC TOXICITY TESTING

Where the aquatic risk assessment has resulted in an unacceptable risk quotient, and further refinement of the exposure assessment (buffer zones, restrictions on application etc) still results in an unacceptable risk to aquatic organisms, higher tier toxicity testing may be considered.

The following guidance in this regard is provided by EC 2002 (c). The source document should be consulted for a fuller explanation and references.

Defining end-points from mesocosm and microcosm studies

The data from microcosm and mesocosm studies should be used to determine a number of end-points that can then be used further in the risk assessment (eg to derive an ecologically acceptable concentration (EAC).

> 'An ecologically acceptable concentration is defined as being the concentration at or below which no ecologically adverse effects would be expected. Depending on the type of study, this may either be defined directly, e.g. from semi realistic multi-species or field studies, or through the application of appropriate uncertainty factors, e.g., with additional single-species tests. For the relevant taxonomic groups in the study, a no observed effect concentration at the community level ($NOEC_{community}$) should be derived using appropriate statistical techniques (eg principal response curves). In addition, NOECs for populations of relevant organisms should be reported ($NOEC_{population}$). Where there are effects at the community or population level, the time taken for recovery to occur should also be reported.
>
> The $NOEC_{community}$, the $NOEC_{population}$ and the time taken for recovery should then be used to determine a no observed ecologically adverse effect concentration (NOEAEC). The NOEAEC is defined as being the concentration at or below which no long-lasting adverse effects were observed in a particular higher-tier study (eg mesocosm). No long-lasting effects are defined as those effects on individuals that have no or only transient effects on populations and communities and are considered of minor ecological relevance – for example, effects that are not shown to have long-term effects on population growth, taking into account the life-history characteristics of the organisms concerned. Different recovery rates may therefore be acceptable for different types of organisms.
>
> The NOEAEC can therefore be higher than the $NOEC_{community}$ or $NOEC_{population}$. Thus, if at a single test concentration effects were determined but recovery occurs and the effect is considered of no concern for the ecosystem sustainability, that concentration should be used as NOEAEC. Different NOEAECs may be derived from a study depending on the protection aim (e.g. in-crop versus off-crop area). When a NOEAEC is derived for a particular study, all of the NOECs that are lower than the NOEAEC must also be presented in order to facilitate interpretation. The lack of ecological relevance of these NOECs must also be justified. The NOEAEC may be used for a direct comparison with the relevant PEC if uncertainty has been reduced considerably and the result from the study is relevant for overall decision making. However, this will require clear knowledge of all relevant end-points and long-term effects. Otherwise an appropriate uncertainty factor should be applied leading to the EAC. Expert judgement is needed in the derivation of an EAC.'.

While the NOEAEC is study specific, the EAC is derived from an overall evaluation of a compound. In concept it is comparable to the predicted no effect concentration (PNEC) defined for other chemical types in the EU framework (eg industrial chemicals, biocides, veterinary medicines, feed additives). However, there is not too much experience with the use of the PNEC in higher-tier risk assessments and clearer differences might emerge in future. Therefore, both terms should be used in parallel for the time being.

Microcosm

Studies with realistic exposure conditions

The environmental fate properties of a pesticide can be an important factor in the mitigation of risk under realistic environmental conditions. If dissipation is rapid, risk assessments based on toxicity studies performed under constant exposure conditions may overestimate potential risks. It is possible to simulate such fate dynamics experimentally in higher-tier studies. Initial indications of the potential influence of exposure on toxicity may be derived for some chemicals (principally those that readily hydrolyse or substantially adsorb) by comparing the results of static and flow through toxicity tests for the same end-point. If apparent toxicity is significantly less in static tests, then fate processes may significantly mitigate risks under natural conditions.

Modified exposure studies are appropriate to address both acute and chronic concerns. One approach to modified exposure studies is to alter the test system to allow a certain environmental fate process to take place, for example, by the addition of sediment to the test system to simulate adsorption or degradation, or by exposing

the test system to natural light conditions to simulate photolysis. In fate simulation studies the method used should be justified on the basis of its relevance to realistic environmental conditions.

Currently, no test guidelines are available for testing algae, *Daphnia*, fish and aquatic plants in water-sediment systems. However, there is some experience with tests on sediment-dwelling organisms. In general, the test organisms should be inserted before the test substance is applied. The test material should usually be applied to the water column of the water-sediment system, but other types of exposure might be reasonable for special purposes. Deviating exposure regimes should be used with care because data may only be related to a single use situation. In such cases, the protocol should be discussed with the regulator or assessing agency.

It is often advisable to determine sub-lethal end-points, even if only acute exposure is expected to reduce uncertainty for such critical substances. Even a short-term exposure may lead to sub lethal effects and this kind of uncertainty is especially relevant in connection with these types of higher-tier studies. The influence and sensitivity of parameters such as the composition of the sediment, the sediment to water ratio, suitable organic carbon content of the sediment, sediment depth, optimal performance of standard species in such tests systems are not yet well understood. Further work is needed to develop specific testing conditions that, on the one hand, are representative of environmental conditions and, on the other hand, ensure that the potential risk is not underestimated.

Microcosm - indoor multi-species tests

Indoor semi-realistic microcosms tests are experiments in systems that intend to represent natural assemblages of organisms characterised by several trophic levels and that, at least for the larger part, are constructed directly with samples of natural ecosystems. Species covering a wide range of sensitivities and biological diversity can be included. In general, indoor semi-realistic microcosms can include micro-organisms, planktonic, periphytic and benthic algae, zooplankton, meiofauna, macroinvertebrates and, when large enough, also macrophytes. Many of the fundamental issues relating to semi-realistic laboratory microcosms also apply to equivalent outdoor studies in mesocosms. There are several advantages of indoor semi-realistic laboratory microcosm tests over outdoor field tests:

- They can usually be run throughout the year. However, since they are constructed in part with samples from natural ecosystems, they closely depend on seasonal availability of biological material.
- There may be potential for a higher level of control over the experimental conditions when compared with an equivalent field system.
- Compared with outdoor studies, set-up costs can be less for tests in laboratory microcosms. However, for a given study design, costs for biological and chemical analysis are similar to outdoor studies.

There are a number of potential disadvantages of semi-realistic laboratory microcosms over larger outdoor mesocosms that should be considered in the selection of an appropriate risk assessment tool:

- They do not usually allow realistic population densities of large organisms (eg fish, newts, frogs and nymphs of larger insects). What is more, if these animals are allowed to be present in indoor semi-realistic laboratory microcosms, they can unduly disturb the test system.
- Long-term effects and recovery of species with complex life cycles may be difficult to determine in indoor test systems.
- There is a lower level of field realism compared to outdoor tests because natural fluctuations in climatic conditions usually are not covered (although these can be simulated).
- The number of micro-habitats present in indoor test systems is usually limited.
- Adequate sampling without overly disturbing certain populations (eg macro invertebrates and macrophytes) can be problematic.

Free living macro invertebrates, however, may be sampled by means of artificial substrates, identified alive, and returned again in the test system. The biomass of rooted macrophytes and the abundance of sediment-dwelling macro-invertebrates usually can be assessed in an adequate way at the end of the experiment only. An alternative approach might be the use of *in situ* bioassays with representatives of these organisms that can be sampled more frequently.

Mesocosms offer the same advantages as microcosms, but in addition, they usually include a wider range of species and generally offer a greater potential to assess the response at the population and, especially, the community level. Furthermore, natural fluctuations in climatic conditions enhance the level of field realism. In particular, they enhance the probability of recovery of some species through eg colonisation. Clearly, the individual concerns arising from the use of a substance must be investigated, and the test design must be tailored accordingly (on a statistically sound basis). However, there is also an argument for some standardisation of a microcosm and mesocosm study design in order to make data for different substances more comparable and ease the interpretation of results. Mesocosm studies are useful in risk assessments when laboratory studies (lower- and higher-tier) indicate potential risks and they should be designed to test specific hypotheses about ecological effects. Mesocosm studies should focus on population level and community-level effects in order to derive an NOEAEC.

Further considerations

An exposure-response experimental design with replication is clearly preferred to ease data interpretation. If possible, this should include the maximum PEC. The selected concentrations should generally be based on the expected effects and not only on the PEC. Previously studies have generally attempted to simulate field exposure (simulation approach). Studies where the chemical is uniformly dosed into the water (toxicological approach) are preferred. They are often more easily interpreted and can be extrapolated to a variety of risk assessment scenarios. Application of the test substance should be made in the period between spring and midsummer when the communities are in their growth phases. Within this timeframe, species richness and abundance are usually most suitable, and the potential time available to observe rates of recovery is long. Due to the density dependence of numerous ecological phenomena, the evolution of small and large systems will be different. For example, the species richness is frequently positively correlated with the size of an experimental system. Due to the relation between functional and structural properties of communities and food webs and the size of the system, the response of a mesocosm to the contamination by a toxicant is not independent of its size. Self-sustainability of the test systems should be taken into account. In particular, the size/complexity of the experimental system should be sufficient to:

- ensure the development and reproduction of the organisms which are being studied
- give sufficient refuges to prey to avoid elimination by predators
- make the recycling of nutrients possible
- ensure potential functional redundancy.

The possibility of recovery may also depend on the size of the systems since large systems may be more resilient than smaller ones to toxicant effects. In general, when constructing a mesocosm, efforts should be made to introduce all the functional groups. This includes primary producers and the various levels of consumers, avoiding introduction of top predators that may greatly influence the system.

Studying fish in mesocosms can present difficulties and needs to be carefully considered. When the invertebrate community is the principal end-point of the study, it is recommended that free-living fish are not included. Macrophytes are an important structural and functional component of shallow aquatic ecosystems, and in general should be included in micro and mesocosm studies that aim to simulate these environments. If macrophyte communities are to be the principal end-point of the study, special efforts are required to establish a diverse and representative community. Efforts should be made to emphasise the use of macrophyte species with relatively low growth potential, otherwise an experimental system might be deeply altered in their response to contaminants.

The notifier should indicate the precise location of the experimental units, and information should be given on the respective location of control and treated systems. The presence of neighbouring natural ecosystems in the immediate vicinity of the experimental area should also be roughly indicated, if it influences the potential for recolonisation of the mesocosm. The level of identification should be as high as scientifically justified or practically feasible (recognising that there are constraints on species identification, especially for smaller species). Special efforts should be made for those groups that are identified in lower-tier studies as potentially the most sensitive. Univariate statistical methods are recommended for investigating effects at the population level, and multivariate methods are recommended for describing community-level effects. The principal response curve (PRC) method is a suitable multivariate technique designed to analyse microcosm and mesocosm tests.

The statistical treatment of data is very important and the use of the aforementioned multivariate technique is recommended to gain insight into the often complex changes in community structure over time and the possible

relationship with treatment. However, the outcome of such evaluations should be carefully checked in the light of the raw data especially for the most sensitive end-points.

Evaluation of test results

When reviewing the results of mesocosm studies, all groups and species should generally be considered of equal importance, as it is difficult to identify the key species. Structural and functional end-points are in general of the same importance. Species structure is usually the principal end-point. Functional end-points alone are not considered appropriate for protecting biodiversity that is the most important assessment end-point. Therefore, in general, differences in species composition at the end of the study between treated test units and untreated controls represent an effect unless these differences can be explained in terms of natural or incidental variations in population and community development. It is important that a sufficient number of populations were present in the study to reach valid conclusions with respect to the most relevant uncertainty factor. Usually there are a few species available with high abundances for which univariate statistical methods can be used. A second group of species occurs usually with lower abundances but the data for the controls give a conclusive picture on the occurrence of these species in the study. Furthermore, there should be a tendency of increasing effects with higher concentrations or clearly no effects in all treatments. These species are also important and the data can be evaluated with multivariate techniques. They are also relevant for a decision upon the uncertainty factor. However, there is a third group of species that are scattered about controls and treatments randomly with highly diverging abundances. These species are usually not relevant for the decision on the most appropriate uncertainty factor. It is particularly important to consider those groups of organisms that were identified as the most sensitive in the standard risk assessment.

For certain taxa or end-points, effects observed in a field study may be considered acceptable, if with appropriate expert ecological judgement, it were considered that they would not pose significant ecological risks to natural aquatic ecosystems. In general, to demonstrate an acceptable level of effect from a particular treatment regime there must be evidence that the treated system and controls are in a comparable state at the end of the study. Test duration should be long enough to be able to observe recovery. For a rough orientation – and to facilitate communication in workgroups - on the overall level of concern related to aquatic ecotoxicology, the following guidance for assessment of effects can be used:

Class 1: effect could not be demonstrated

- no (statistically significant) effects observed as result of the treatment, and
- observed differences between treatment and controls show no clear causal relationship.

Class 2: slight effect

- effects reported in terms of slight or transient and/or other similar descriptions, and
- short-term and/or quantitatively restricted response of sensitive end-points, and
- effects only observed at individual samplings.

Class 3: pronounced short-term effect

- clear response of sensitive end-points, but total recovery within 8 weeks after the last application, and
- effects reported as temporary effects on several sensitive species, temporary elimination of sensitive species, temporary effects on less sensitive species/end-points and/or other similar descriptions, and
- effects observed at some subsequent sampling instances.

Class 4: pronounced effect in short-term study

- clear effects (such as strong reductions in densities of sensitive species) observed, but the study is too short to demonstrate complete recovery within eight weeks after the (last) application.

Class 5: pronounced long-term effect

- clear response of sensitive end-points and recovery time of sensitive end-points is longer than 8 weeks after the last application, and
- effects reported as long-term effects on many sensitive species/end-points, elimination of sensitive species, effects on less sensitive species/end-points and/or other similar descriptions, and
- effects observed at various subsequent samplings.

The following suggestions about the translation of effect classes into NOECs and NOEAECs may be considered. If only effects related to Class 1 were observed, the NOEC and NOEAEC are the same, which is not the case for effects belonging to the other classes.

With respect to Class 2 effects, a NOEC and a NOEAEC should be determined although the values should often be the same. There is a need to explain that effects occurred, but that these effects were regarded for some reasons as ecologically not adverse.

For effects in Class 3, a clear difference between the NOEC and NOEAEC should be determined.

A NOEAEC cannot be determined if effects belonging to Class 4 and 5 were observed. Whilst for Class 4 effects, it may be possible to use other tools (see below), to show that effects are acceptable; this could be very difficult for effects belonging to Class 5. Intrinsic recovery potential mainly relies on resting stages present in the treated system itself (eg resting eggs of Cladocera or rotifers, algal spores). The importance of this phenomenon will frequently be dependent on the duration of the pre-exposure period since resting stages are naturally produced when climatic conditions become unfavourable. Therefore, if the systems experienced one or more autumn-winter-spring cycle before treatment, the abundance would be greater than for recently built mesocosm.

The notifier should, therefore, indicate the precise history of the systems. Effects may be considered of low ecological significance if recovery takes place in a given time period like eight weeks, but this period should not be used as strict trigger because recovery depends very much on the life history of the species. Even if recovery is observed in a mesocosm study, the extent and rate of recovery has to be considered in the context of natural aquatic systems and the proximity of unaffected sites to those affected. Where recovery of a species is not observed, or is only incomplete in a mesocosm study, it is the responsibility of the data submitter to discuss this observation and explain how this relates to the likelihood of recovery in natural aquatic ecosystems.

Furthermore, some species cannot recover in mesocosm studies simply because of the conservative study design (eg gammarids). It is recommended that additional tools (eg.further laboratory studies) are used to address the remaining uncertainty. The replacement of species is not acceptable in general. But in some cases, the replacement of one species by another with a similar role in the ecosystem may be considered acceptable (eg for some algal species) if functionality is maintained and no further structural effects occur (eg no indirect effects on zooplankton). The replacement species, however, should have a similar function. For example, a replacement of green algae by blue-green algae or photosynthetic-facultative flagellates is unacceptable. In any case, functional characterisation of mesocosms should be performed for a significant period of time since functionality may sometimes be maintained for a short-term period but may decrease later. The notifier has to provide clear evidence that the ecological function and community structure in the field situation is unlikely to be significantly affected.

It is recommended that for all species affected in a mesocosm study, the likelihood of recovery under field conditions be fully addressed when evaluating the study results. All factors that may influence population/community recovery should be considered, and should include dispersal ability, life-history, breeding season, number of breeding attempts per season, abundance in the environment, spatial records, as well as the natural variability in population sizes and distributions. Population-level evaluation of genetic properties should also be considered. Genetic variability is a matter of concern since spatially limited populations that develop in mesocosms may exhibit significant differences in various characteristics (eg consanguinity, founding effects) as compared to natural populations of the same species. If the same experimental systems are used from one study to another, the case of selection of less sensitive genotypes cannot be excluded. In this case the evaluation of effects may be biased (underestimation of effects). Increased homozygoty may also alter the pattern of response of some species to pesticides. It is therefore recommended to replace sediment after an experiment before a new mesocosm study is started in the same testing facility. Results of field studies should be accompanied by clear explanations as to why a given observed effect should be considered ecologically significant or acceptable when they are presented to regulatory authorities and that, wherever possible, such studies should be reviewed by groups of experts to provide the least-biased advice, although it is accepted that this may be difficult under current registration procedures. Connections could also be established with experts working in the field of biological conservation.