National Institute
for Public Health
and the Environment

Report 601712004/2008

B.J.W.G. Mensink | C.E. Smit | M.H.M.M. Montforts

Manual for summarising and evaluating environmental aspects of plant protection products

rivm

RIVM Report 601712004/2008

Manual for summarising and evaluating environmental aspects of plant protection products

B.J.W.G. Mensink
C.E. Smit
M.H.M.M. Montforts

Contact:
C.E. Smit
Expertise Centre for Substances
ce.smit@rivm.nl

This investigation has been performed by order and for the account of Netherlands Ministry of Housing, Spatial Planning and the Environment, within the framework of project M/601712, 'Consultancy on pesticides and biocides'.

Dedicated to the memory of Hans Mensink

† 21 May 2008

rivm

Rapport in het kort

Handleiding voor het samenvatten en evalueren van milieuaspecten van gewasbeschermingsmiddelen

Het RIVM heeft een handleiding opgesteld om studies voor het beoordelen van milieurisico's van gewasbeschermingsmiddelen samen te vatten. De handleiding geeft vervolgens aanwijzingen om de betrouwbaarheid te kunnen beoordelen van verschillende soorten (standaard)toetsen over de afbraak, verspreiding en effecten van deze stoffen in het milieu. Ook geeft het rapport aan hoe het onderzoeksresultaat voor de uiteindelijke risicobeoordeling kan worden gebruikt. Deze handleiding vergroot de eenduidigheid tussen beoordelingen van verschillende organisaties.

De eerste versie van deze handleiding heeft het RIVM in 1995 uitgebracht. Nadien is de inhoud in interne tussenrapportages verschillende malen aangepast. Met het huidige rapport is wederom een openbare versie beschikbaar.

Trefwoorden: handleiding; dossier evaluatie; toelating; milieu; gewasbeschermingsmiddelen

rivm

Abstract

Manual for summarising and evaluating environmental aspects of plant protection products

This RIVM report is a guidance document for summarising and evaluating study reports that are used for the environmental risk assessment of plant protection products. For different types of (standard) tests in the field of fate and behaviour and ecotoxicology, guidance is given to assess the scientific reliability of the studies and to determine the usefulness for risk assessment.

This Manual was first published in 1995 as an RIVM report. Since then, several revisions were issued in the form of internal quality documents. With the present report, a new update is publicly available to promote the transparency and uniformity of the evaluation procedure.

Key words: guidance; dossier evaluation; registration; environment; plant protection products

rivm

Preface

This report is based on the 'Manual for summarising and evaluating the environmental aspects of pesticides' by Mensink et al. (1995). After the first publication in 1995, the contents have been revised and updated several times. Being internal quality documents, the updates have never been publicly available. In the Dutch process of (re-)registration of plant protection products and biocides, several scientific institutions and consultants take part in the evaluation process, and it has been considered as a drawback that the updates by RIVM were not available to the other evaluating institutes. The current update of the 'Manual' is again issued as a publicly available RIVM report, to promote the transparency and uniformity of the evaluation procedure. This version should be regarded as an interim version, that will be open for comments from (inter)national experts and will be revised and extended with new information in the foreseeable future. Any comments are thus highly appreciated and can be sent to the corresponding author via ce.smit@rivm.nl

rivm

Acknowledgements

Thanks are due to Dr Hans Vonk (EPP Consultancy) and Dr Frank de Jong (RIVM) for their contributions to the report and/or comments on earlier drafts.

rivm

Contents

rivm

List of tables and figures

Tables

Figures

rivm

List of terms and abbreviations

a.i.	Active ingredient
AR	Applied radioactivity
BBA	Biologische Bundesanstalt
BCF	Bioconcentration Factor = ratio between concentration in (parts of) biota and surrounding medium at steady state
BR	Bound Residues
CEC	Cation Exchange Capacity
Ctgb	College voor de Toelating van Gewasbeschermingsmiddelen en Biociden; Dutch Board for the Authorisation of Plant Protection Products and Biocides
CV	Coefficient of Variation
DOC	Dissolved Organic Carbon
DT_{50}	Degradation half-life = time in which 50 % of the parent compound has degraded
EC	European Commission
EC_{50}	Median Effective Concentration = concentration expected to cause a 50 % change of a parameter relative to the control (note: in *Daphnia* studies, the EC_{50} is the concentration at which 50 % of the tested population experiences sub-lethal effects as compared to the control)
EPA	Environmental Protection Agency
EPPO	European and Mediterranean Plant Protection Organization
EU	European Union
fd	Feed
FOCUS	FOrum for the Co-ordination of pesticide fate models and their USe
GLP	Good Laboratory Practice
HTB	Handleiding voor de Toelating van Bestrijdingsmiddelen
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
K_d	Distribution constant
K_F	Freundlich adsorption coefficient
K_{om}	Sorption coefficient normalised to organic matter
K_{ow}	Octanol-water partition coefficient
$K_{s/l}$	Sorption coefficient or soil-water partition coefficient
LC_{50}	Median Lethal Concentration = concentration expected to kill 50 % of the exposed organisms
LD_{50}	Median Lethal Dose = dose expected to kill 50 % of the dosed animals
LR_{50}	Median Lethal Rate = application rate expected to kill 50 % of the exposed organisms
MHWC	Maximum Water Holding Capacity
NOEC	No Observed Effect Concentration = highest tested concentration at which the test parameter shows no significant difference as compared to the control
NoRC	No Repellent Concentration = concentration at which no repellency is observed
OECD	Organization for Economic Cooperation and Development
OM	Organic matter
PEC	Predicted Environmental Concentration
pF	Suction tension in soil
pH	Measure of the acidity or alkalinity of a solution
pK_a	-log K_a. K_a = dissociation constant
PPO	Institute of Applied Plant Research (Wageningen University and Research Centre)
Rf	Retardation factor

Ri	Reliability indicator
RIVM	Rijksinstituut voor Volksgezondheid en Milieu; National Institute for Public Health and the Environment
RIVM-SEC	Expertise Centre for Substances - RIVM
SI	Système International d'Unités
Soil TLC	Soil Thin/Thick Layer Chromatography
STP	Sewage Treatment Plant
TGD	Technical Guidance Document
ThOD	Theoretical Oxygen Demand
USDA	United States Department of Agriculture
ww	Wet weight
fw	Fresh weight

rivm

Samenvatting

De toelating van gewasbeschermingsmiddelen en biociden is gebaseerd op dossiers die worden ingediend bij de Competente Autoriteit, in Nederland het College voor de Toelating van Gewasbeschermingsmiddelen en Biociden (Ctgb). De dossiers bevatten rapporten van studies die ingaan op de mogelijke risico's voor mens en milieu. Deze rapporten moeten worden samengevat en geëvalueerd om de eindpunten af te leiden die in de risicobeoordeling kunnen worden gebruikt. Het Ctgb heeft deze taak uitbesteed aan verschillende instituten. Deze handleiding heeft tot doel de eenduidigheid tussen die instituten te bevorderen, daarmee de inzichtelijkheid in het beoordelingsproces te vergroten en het vastleggen van de uitgangspunten voor de risicobeoordeling te vergemakkelijken.
De handleiding richt zich voornamelijk op studies naar het gedrag in bodem en water en de ecotoxicologie. Er worden richtlijnen gegeven voor het beoordelen van de betrouwbaarheid van individuele studies en voor het opstellen van studiesamenvattingen. Ook wordt aandacht gegeven aan de bruikbaarheid van gegevens voor de risicobeoordeling. Hoewel in eerste instantie opgesteld in de context van de toelating van gewasbeschermingsmiddelen en biociden, kan de handleiding ook worden gebruikt voor de evaluatie van studies binnen andere kaders.

rivm

Summary

The (re-)registration of plant protection products and biocides is based on dossiers that are submitted to the Competent Authorities, which in The Netherlands is the Dutch Board for the Authorisation of Plant Protection Products and Biocides (Ctgb). The dossiers contain test reports that address the potential risks to humans and/or the environment. The reports have to be evaluated to derive the endpoints that are needed for risk assessment and a summary of methods and results has to be prepared. Ctgb has commissioned this task to different institutes. This manual aims to promote the uniformity among the evaluating institutes and thereby the transparency of the evaluation process, and to facilitate the proper documentation of the decisions underlying the risk assessment.
The Manual focuses primarily on studies on fate and behaviour in soil and water and ecotoxicology. Instructions are given on the evaluation of the intrinsic reliability of the studies and on the preparation of study summaries, and the usefulness of results for risk assessment is addressed. Although originally developed within the context of registration of plant protection products and biocides, the guidance can be used for study evaluation in other frameworks as well.

1 Introduction

This present report is an updated version of the report of Mensink et al. (1995), which was prepared to provide guidance for summarising and evaluating the environmental aspects of plant protection products and biocides for regulatory purposes. For the (re-)registration of these compounds, the applicant submits a dossier to the regulatory authorities, containing test reports that are needed to evaluate the potential risks to humans and/or the environment. The reports have to be evaluated to derive the endpoints that are needed for risk assessment and a summary of methods and results has to be prepared. The purpose of this report is to improve:

- the reproducibility of the evaluation process;
- the traceability of the evaluation process;
- the proper documentation of the decisions underlying the risk assessment.

Instructions are given on the evaluation of the intrinsic reliability of the studies and on the preparation of study summaries, and the usefulness of results for risk assessment is addressed. The guidance in this report reflects the way evaluation reports are prepared at the Expertise Centre for Substances of the National Institute for Human Health and the Environment (RIVM-SEC) within the context of the national registration procedure in The Netherlands. The guidance thus focuses on the preparation of advisory reports for the Dutch Board for the Authorisation of Plant Protection Products and Biocides (College voor de Toelating van Gewasbeschermingsmiddelen en Biociden, Ctgb), but the information in this report is considered valuable for study evaluation in other frameworks as well.
Extensive information on the registration procedure can be found in the Handbook for the Registration of Pesticides (Handleiding voor de Toelating van Bestrijdingsmiddelen, HTB), which is available at http://www.ctgb.nl/.

2 Data quality, reliability and usefulness

2.1 Reliability indicators

The reliability pertains to the intrinsic, scientific value of an individual study. The reliability is determined by the set-up, performance and evaluation of the experiment, and the reporting. A properly reported study may be considered less or not reliable because of an inadequate set-up (e.g. too few replicates), performance (e.g. control mortality too high) or data evaluation (e.g. inadequate statistics). Likewise, a study that was originally carried out in a scientifically sound way, may be classified less or not reliable in case the description is very concise (e.g. experimental set-up is given as a reference to another report), or if various items that are considered important for interpretation of the test results cannot be checked (e.g. temperature data are not given).

Reliability indicators (Ri) are used to designate the reliability of a test or study. Ri is 1, 2, or 3 reflecting reliable, less reliable, and unreliable test results. For an explanation of the criteria see Table 2-1[1].

Table 2-1 Reliability indicators for qualifying public literature and confidential scientific reports from industries

RELIABILITY INDICATOR	DESCRIPTION	DEFINITION
Ri 1	RELIABLE	the methodology and the description are in accordance with internationally accepted test guidelines and/or the instructions in this manual
Ri 2	LESS RELIABLE	the methodology and/or the description are less in accordance with internationally accepted test guidelines and/or the instructions in this manual
Ri 3	NOT RELIABLE	the methodology and/or description are not in accordance with internationally accepted test guidelines and/or the instructions in this manual

From the above table it appears that conformity with accepted guidelines is a key element for the assignment of a reliability score. Various (inter)national organisations are active in the field of test guideline development. Nowadays, tests submitted for regulatory purposes will most often be performed according to the OECD Guidelines for the Testing of Chemicals, which include most relevant internationally agreed test methods used by government, industry and independent laboratories. Older studies, and studies into aspects that the OECD Guidelines do not cover, may be performed according to other guidelines. Links to information on guidelines are provided in Annex 3.
In principle, only Ri 1 and Ri 2 studies are used for environmental risk assessment of pesticides. Therefore, it is particularly crucial to have clear criteria to discern between Ri 2 and Ri 3 studies, as Ri 3 studies are not used for environmental risk assessment. However, this is not only a matter of using

[1] Similar schemes for assigning reliability indices to studies have been implemented in other frameworks, see e.g. the Technical Notes for Guidance on Dossier Preparation including preparation and evaluation of study summaries under EC Directive 98/8/EG (biocides directive). Although based on the same principles, these schemes do not necessarily lead to the same reliability indicator.

criteria as transparent as possible. It is a matter of reviewer experience as well, specially when studies have to be evaluated that were not performed to fulfil data requirements for (re)registration in particular. It is important to note that the scientific reliability of any individual test is always determined in the context of the data requirements for (re-)registration. This implies that the methodology may be insufficient or insufficiently reported to extract reliable data for risk assessment, whereas the study in itself may have been a reliable study to test the hypothesis of the author self. The reviewer should be well aware of this.

2.2 Usefulness

The usefulness indicates whether a study is appropriate for a particular purpose, e.g. environmental risk assessment for the authorisation of pesticides or for standard setting procedures. Where reliability generally refers to an individual study, usefulness thus refers much more to a study in relation to other comparable studies, and to the choice which study or studies match the best with a particular purpose. Scientific reliability is a prerequisite for a test to be used for registration purposes. The next step is to decide whether a valid endpoint (i.e. reliable or less reliable, Ri 1 or 2, not unreliable, Ri 3) can be used in environmental risk assessment. Two main questions can be identified:

1. Does the test deliver the endpoint that is needed for risk assessment?
2. Is the test representative for the situation that is to be assessed?

The first question may seem obvious, but sometimes tests are supplied that are highly reliable in a scientific way, but do not deliver the endpoint that is needed for risk assessment (e.g. a 7-days earthworm test, whereas the trigger is based on a 14-days study). The second question refers to the relevance of the test for risk assessment items in terms of test conditions, test medium, and type, time and place of application. As an example, laboratory data on soil degradation performed with volcanic or paddy soils, or with concentrations deviating from those expected result from the use of the compound, may be not useful. Similarly, a study of the biodegradability of a pesticide under arctic conditions is not very useful to describe the dissipation route that is likely to occur in the Netherlands. Field data obtained in crops or climates different from those under consideration, may also be not useful. This manual focuses primarily on the evaluation of the reliability, but in some cases guidance on the determination of the usefulness for risk assessment is given as well.

2.3 Selection of endpoints for risk assessment

As was stated above, only data with Ri 1 or Ri 2 can be used for the risk assessment, provided that they are useful (see Figure 2-1). In general, both Ri 1 and Ri 2 data — as equally useful — are included in the final calculations. If there are enough Ri 1 data, however, it may not be necessary to include Ri 2 data. In case Ri 2 data are very different from Ri 1 data (e.g. Ri 1 data indicate toxicity, whereas Ri 2 data indicate slight toxicity), the final choice may be dependent on the weight of evidence. If Ri 2 data are not selected for risk assessment, this should be motivated.

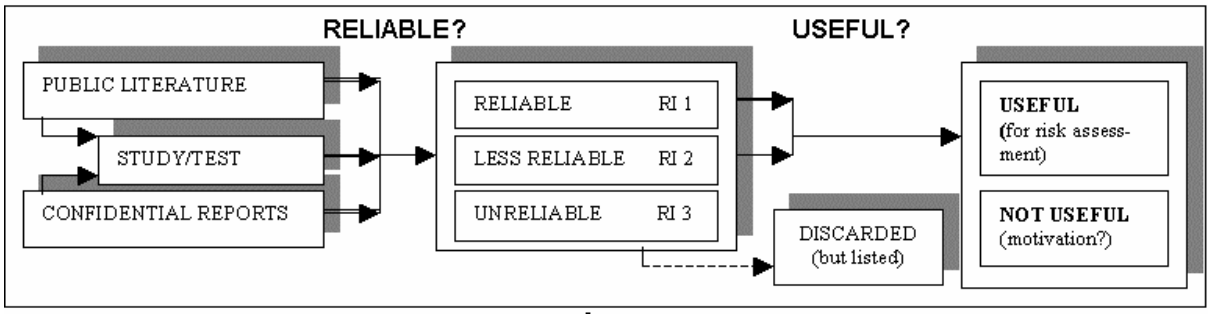


Figure 2-1 Selection of test results depending on reliability and usefulness

2.4 Quality of public versus confidential sources

Regulatory authorities have to deal with different kinds of information sources. The two main categories of such sources are the unpublished, confidential reports from chemical industries — generally submitted for registration purposes —, and the public literature in scientific journals. The former will prevail when a new pesticide has to be marketed, of which the research and development has been a matter of involvement of a small group of scientists, belonging to or contracted by a company. They are generated by scientific laboratories and submitted to the regulatory authorities on a confidential base. These unpublished reports generally comply with the current data requirements of the regulatory authorities for (pesticide) registration. They are generally performed under GLP (Good Laboratory Practice) and in accordance with recent guidelines, and include raw data. Therefore, they are suited for verification.

Public literature — i.e. published scientific articles and reports that are publicly available —may prevail, when a pesticide, already on the market, has been investigated by scientists, for whatever reason (e.g. as a model substance to investigate particular ecological processes). This public literature does not necessarily comply with the data requirements for pesticide registration. The studies are also not necessarily in accordance with GLP, and raw data are generally not available. This means that for the evaluation of pesticides within the context of the Dutch authorisation procedure, these studies will at most receive an Ri 2 and will in most cases not be used because better data are available. Within other frameworks, where dossiers merely consist of open literature, it may be decided to use them as acceptable or additional information.

3 How to summarise and evaluate studies?

3.1 Study summaries

Study summaries form the basis of evaluation reports. A summary is a concise text containing the most relevant information that is necessary for the interpretation of the study and its results, and the decision on the use for environmental hazard and risk assessment. A summary contains a *Header*, which is a table listing the most relevant test conditions, the endpoints that are derived from the study and an indicator of the reliability of the test. Unreliable data are not listed in the header, but the experimental conditions are included to show that the experiment was conducted (see Table 3-1).

Table 3-1 Example of a header

Substance	Soil type	OM [%]	pH	pF	Condition	Dose [kg as/ha]	Dose [mg as/kg][1]	T [° C]	Duration [d]	DT_{50} [d]	DT_{50} 20 °C [d]	Ri
^{14}C-XXX	loam	3.5	5.1	2	aerobic	1 x 0.84	1.1	25	329	167	249	1
^{14}C-XXX 25 % EC	loam	3.5	5.1	2	aerobic	1 x 0.84	1.1	25	329	144	215	1
^{14}C-XXX	loam	3.5	5.1	2	aerobic	1 x 8.4	11	25	329			3
^{14}C-XXX 25 % EC	loam	3.5	5.1	2	aerobic	1 x 8.4	11	25	329	431	643	2

1: assuming 5 cm soil depth and soil bulk density 1500 kg/m^3

The summary continues with a *Description* of the methodology followed by the *Results*, and ends with a *Remarks* section in which critical comments or recalculations of the reviewer are reported. Depending on the purpose of the advisory report, the Description is a text block (e.g. for EU monographs) or may be structured as list of items (national evaluations). For complex field studies, it may be appropriate to adapt the structure of the summary, see e.g. De Jong et al. (2008); De Jong et al. (2006), who recommend to include an *Abstract* (containing the decision making information), *Extended Summary* (a summary of the study report), *Evaluation* (critical comments on the test, evaluation of reliability and results) and *Suggestions for the use in risk assessment.*

3.2 Summary Tables

Study summaries are prepared on the basis of the so-called Summary Tables that are the core of this document. For a large number of test types, they list all relevant aspects of environmental test items that may contribute to the test's reliability and usefulness. Summary Tables comprise a wide array of test items that may influence the test quality, they function as a checklist and they refer to scientific guidelines, when necessary. The guidance in this report generally coincides with internationally agreed protocols and guidelines, e.g. from the OECD, (US)EPA or EU. However, the instructions in this report pretend to summarise these protocols and guidelines in a systematic way, and are more profound when necessary, particularly in that additional information is given on e.g. how to (re)calculate certain endpoints. The tables with their test items may help to structure the abundance of information and to tag an Ri to a particular test or a part of it. Summary Tables can always be extended and specified, concurrent with scientific developments. In this way, the instructions provide an overall checklist for establishing the scientific status of any environmental study and to (re)calculate all data necessary for the risk assessment.
All Summary Tables are structured as presented in Table 3-2.

Table 3-2 Structure of a Summary Table.

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
These items should always be included in the test description of a summary. **If not reported, the reliability is almost always decreased.**	These notes explain the requirements which have to be met for a reliable test (i.e. with an adequate methodology and description). If items in a study deviate from these requirements, check in the next column ('reliability lower?') whether the reliability with respect to that particular item may decrease. Y(es) This note indicates that the reliability can be considered to decrease. E(xpert judgement) This note indicates that no clear guidance can be given here. The assignment of an Ri is up to the reviewer.	 Y E
RESULTS		
These results should always be included, under Results.	X in this column indicates a cut-off value. When there is no consensus, expert judgement should be used to establish one. Cut-off values that are used in the Dutch effect and risk assessment are given here between square brackets (e.g. the application rate should not exceed X times the recommended rate [X=2]).	
PAY ATTENTION TO		
The items here should not necessarily be included in a Summary, but should be checked. These items —if deviating from the requirements— can be included under Remarks.		

If items reported are less or not in accordance with the Summary Tables, the reliability of a study is expected to decrease. In the column with the heading 'Reliability lower?' this is indicated by a Y(es) or an E(xpert judgement):

> **Y**(es) indicates that solely based on not fulfilling this requirement for this item, the reliability of the test as a whole is expected to decrease.
>
> **E**(xpert judgement), indicates that no clear guidance can be given. The reviewer can consult a specialist. The reviewer should clarify the reasons for a decision on tagging a Ri.

When the reliability of a test is expected to decrease, this can be reflected in tagging a Ri 2 to a test, or even tagging a Ri 3. It is up to expert judgement to decide how many 'Y'-items are required for tagging a Ri 2 or 3 to a particular test in its entirety. However, there is a core set of test items that *must* comply with the Summary Tables. If a test does not comply with this core set, the test is considered unreliable and tagged with Ri 3. These items are indicated with [➔ Ri 3]. The core set reflects the following questions:

- Which substance is tested, for how long, and what are its concentrations or application rates?
- Is the test system (abiotic and biotic) clearly reported?
- Is there a valid control, and are the most relevant test conditions clearly reported?
- Is the endpoint clearly reported?

In recapitulation, decisions on the scientific quality of a test should be based on the Summary Tables: a checklist for all relevant scientific aspects of a test. Their interpretation should not be too rigid: expert judgement, reflecting e.g. the weight of evidence, can overrule the instructions in a Summary Table. However, the quality of a test must be in compliance with a core set of test items. If not, the test is

unreliable and should not be used for further risk assessment. When an Ri 2 or 3 has been tagged to a test or study, the reviewer should report clearly, why this has been done.

Specific guidance for individual aspects is given in Chapter 4 and following.

3.3 General issues on endpoint derivation

3.3.1 Rounding off

Results have to be rounded off correctly. Rounding off is done for the end-result after all intermediate calculations have been performed. As a general rule, the following applies:

- If necessary, the final digit of a number is rounded up when it is ≥ 5, and rounded down when it is < 5 (e.g. 8.95→9.0; 8.94→8.9). Rounding off numbers is also dependent of the following:
    - Numbers ≥ 10 are given without decimals, except when a scientific notation has been used in which case the figure is rounded off to two significant digits (e.g. 10.12→10; 108.24→108; 1000.586→1001; $3.86\times10^{2}\rightarrow3.9\times10^{2}$);
    - Numbers < 1-9 are rounded off to two significant digits, also when a scientific notation is used (0.375→0.38; 2.5→2.5; 3.75→3.8; $3.86\times10^{-2}\rightarrow3.9\times10^{-2}$)

3.3.2 Units

In principle, units of the S.I. (Système International d'Unités) are used. However, in (inter)national evaluation reports, L is preferred over dm^3, and dosage is expressed as [kg/ha] instead of [g/m^2]. For recalculating other commonly used units, first consult Appendix 1.

3.3.3 Statistics

Correct statistical procedures need to be followed. Where appropriate, suitable methods for evaluating the test results are dealt with in the respective (sub)sections.

Fitting of data to (non-)linear models needs to be performed with at least five data points. As a general rule, the outcome of a fitting procedure is only accepted when $r^2 \geq 0.8$ (EFSA, 2007b). When $r^2 < 0.8$, the result is not considered valid.

4 Identity and physico-chemical properties

4.1 Names, substances and products

All trade names, chemical names (standard: IUPAC), and other names of the active substance (as) are given. Industrial codes may be used instead of the complete chemical names if that is the only information the company has provided. The use of trivial names is encouraged.
Radioactive substances should be recorded as mentioned above. The position of the label should be given in the Description. Mostly mixtures of labelled and unlabelled material are used; these are referred to as labelled. If the composition of a formulated product is not fully known, this is also remarked in the evaluation.
Metabolites are given by name and/or (company) code. A table with codes, (chemical) names and structural formulas (if known) can be included in the report.

4.2 Physico-chemical properties

The next physico-chemical properties are at least required for the risk assessment with current fate models and/or for the interpretation of test results: log K_{ow}, solubility in water, pK_a, and vapour pressure. Henry's Law Constant H can be calculated from the vapour pressure, molar mass and water solubility according to:

$$H = \frac{P \times M}{S}$$ Equation 1

in which:

H	=	Henry's Law Constant [Pa.m^3/mol]]
P	=	vapour pressure [Pa]
M	=	molar mass [g/mol]
S	=	water solubility [g/m^3]

5 Fate and behaviour in soil and (ground)water

5.1 General aspects

For an introduction in the use of soil science for the evaluation of the behaviour of substances in the soil, the reviewer is referred to Van Gestel (1986).

5.1.1 Storage of soil samples

Sampling, handling and storage of soils should be done according to ISO 10386 – 6 (see further OECD 307). Storage conditions of soil, that is not immediately used for transformation studies, should be as follows: in the laboratory at 4 °C for at most three months (to avoid anaerobic conditions); in the open or in a glasshouse under well-drained conditions (to avoid desiccation). The maximum allowable storage time can be estimated as follows:

$$t = 90 \times e^{(-0.1[T-4])}$$ Equation 2

in which:

t = allowable storage time [d]
T = temperature [°C]

5.1.2 Water content of soil

The pF of the soil can be represented as a function of the soil water content. For the estimation of the pF of a given test soil, the reader is referred to the factsheet of Van Vlaardingen and Smit (2008).

5.1.3 Classification of the soil type

The classification of the soil type, given by the authors, should be checked with the American Soil Classification System (USDA, 1951; see Annex 2), and this US-classification is reported in the Header. If verification is not possible, the classification given by the authors is used (if necessary a literal translation from Dutch, French or German). It should clearly be stated in the Description, which classification was used.
The sizes of soil particles are fastened down in the different classification systems. The sizes are however not identical in the different systems. Pay attention to this when classifying a soil according to the USDA system, a procedure to estimate the correct particle size classes is also given in Van Vlaardingen and Smit (2008).

If it concerns soil types not representative for normal agricultural practice, like paddy soil and volcanic soil, this should be clearly stated.

5.1.4 pH measurement

The pH of the soil can be measured in the water phase (pH-H_2O) of the soil, or after a solution of KCl or $CaCl_2$ was added (pH-KCl and pH-$CaCl_2$, respectively). pH-KCl and pH-$CaCl_2$ are always lower than pH-H_2O as more protons in the soil solution can be measured.

It is assumed that the pH-$CaCl_2$ (0.01 M) gives the best estimate for the soil solution, and is therefore the most convenient value with respect to bioavailability for plants. Record in the Header the pH-$CaCl_2$, if available. If not, record the pH-KCl or the pH-H_2O (in this order of preference). It should always be indicated in the Description, which pH is used.

5.2 Aerobic and anaerobic transformation in soil

5.2.1 General aspects

Summary Table 5-1 refers to soil aerobic transformation studies. DT_{50} values should be based on transformation. Transformation means the compound is converted to smaller or larger molecules by biological, microbiological, and/or chemical action. Degradation means the compound is converted to smaller molecules by biological, microbiological, and/or chemical action. Dissipation means that the compound 'disappears': this can be by transformation, volatilisation, leaching, plant uptake, or run-off. Mineralisation means the compound degrades to inorganic compounds (e.g. H_2O, CO_2). Results of sterile incubations are not presented in the Header, but are given in the Results. Unusual results (e.g. degradation) are mentioned in the evaluation.

5.2.2 Microbial activity of the test soil

Fertilisation and measures to improve soil structure, which are common practice for agricultural soils, generally have a positive effect on microbial biomass and activity. According to OECD 107, microbial biomass should be ≥ 1% of the organic C-content. Additional information is retrieved from Anderson and Domsch (1989), who investigated microbial carbon content in 134 agricultural soils in Western and Central Europe (England, The Netherlands, Denmark, Germany and Poland). About half of the locations had a monoculture, on the other half crop-rotation was applied, soils were in use for 7 to 144 years. From the data it appears that 147 to 734 mg C/kg soil can be considered as a normal range for agricultural soil. If biodegradation tests are performed using soils with a microbial biomass outside this range, the resulting DT_{50} values should be judged in the light of other available information. If the biomass of the experimental soil is < 147 mg C/kg soil, the degradation rate for a viable soil may be underestimated, while for soils with > 734 mg C/kg soil, an overestimation is possible. In case the DT_{50} in these soils is in line with other values, the result can be accepted. In other cases, the following applies:

- For soils with biomass < 147 mg C/kg soil:
    - if the DT_{50} is high in comparison with other values (i.e. slow degradation), this may be due to the relatively low microbial activity
    - a low DT_{50} (i.e. fast degradation) in comparison with other values is not expected
- For soils with biomass > 734 mg C/kg soil:
    - if the DT_{50} is low in comparison with other values (i.e. fast degradation), this may be due to the relatively high microbial activity
    - a high DT_{50} (i.e. slow degradation) in comparison with other values is not expected.

In these cases, the reliability of the results should be critically evaluated.

5.2.3 (Re)calculation of DT_{50} values in soil and curve fitting

Detailed guidance for the (re)calculation of DT_{50} values and their use in risk assessment is given in FOCUS (2006) and EFSA (2007b). DT_{50} values are normalised to a reference temperature using the equation:

$$DT_{50T1} = DT_{50T2} \ exp\left(\frac{E_a}{R}\left[\frac{1}{T_1} - \frac{1}{T_2}\right]\right)$$ Equation 3

in which:
DT_{50T1} = half-life at temperatures T_1
DT_{50T2} = half-life at temperatures T_2
k = rate constant [1/d]
A = factor equal to the rate coefficient at infinite temperature [1/d]
E_a = activation energy [kJ/mol]
R = gas constant (0.008314 kJ/K·mol]
T = absolute temperature [K]

The PPR Panel recommends that the median E_a value of 65.4 kJ/mol corresponding to a Q_{10} of 2.58 should replace the default E_a value of 54.0 kJ/mol corresponding to a Q_{10} of 2.2, which has been used before (EFSA, 2007b).

When converting to 20 °C; the formula can be simplified to:

$$DT_{50,20C} = DT_{50,T} \times e^{0.095(T-20)}$$ Equation 4

in which:
T = temperature [°C] at which the study was conducted.

Table 5-1 Soil, aerobic transformation studies (top soil)

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance and position of label	improperly characterised or reported?	Y [→Ri 3]
2. vehicle	reported? amount used?	E
3. test type		
3.1 aerobic	aerobic, non-sterile conditions are required	Y
3.2 sterile	method of sterilisation should be given	Y
4. soil	top soil should be used; no enrichment with e.g. alfalfa	Y
4.1 soil type	USDA-class and other relevant data reported? soil must be relevant for the agricultural situation (e.g. no paddy or volcanic soil; 0.5 – 15 % OM)	Y
4.2 characteristics	pH, CEC, OM reported?	Y
4.3 agricultural history	reported? no prior use of compounds that may have lead to adapted micro-organisms in the previous five years	E
4.4 storage conditions	if there is no immediate use, storage in the lab or in the open should be appropriate (see text) BBA/Speyer soils before 1982 were probably stored too dry	Y
4.5 microbial activity	not determined? microbial biomass of agricultural soil 147 – 734 mg C/kg	E
5. weight of soil sample	weight soil sample should be ≥ X g. [X= 25]	Y
6. temperature	monitored? constant?	Y
7. application	way of application reported?	Y
7.1 rate	within X times [X = 2] the recommended application rate[a]	E
8. moisture content	in an aerobic study: pF 2 – 4[b]	Y
9. light condition	incubation in the dark (unless photodegradation has been shown to be of no importance)	Y
10. test system	should be closed with volatile traps	E
11. incubation time	preferred until 90% transformation or up to 100 days	E
12. sampling frequency	≥ 5 time points are needed for adequate regression analysis	E
13. extraction/analysis method	should be appropriate for the substance and the metabolites, and the recovery of the substance should be >X% [X= 70] and < Y% [Y = 110]	E
RESULTS		
1. distribution of radioactivity over time	present table	
1.1 recovery of radioactivity: total, extractable, bound residues (BR), CO_2, volatiles	total recovery at every time point should be > X% [X = 80] (recovery of radiolabel or the sum of compounds) BR and CO_2: amount after 100 days is endpoint	E
1.2 radioactivity assigned to parent and metabolites	identified and quantified separately?	Y
1.3 relevant metabolites	relevant: metabolites: > 10 % of AR, > 5 % of AR at 2 consecutive endpoints or maximum reached at end	
2. DT_{50} of parent and metabolites	give way of calculation	
PAY ATTENTION TO		
1. dissipation type	this should be transformation	E
2. microbial biomass	≥ 1% of organic C if microbial C is lower than normal range, and DT_{50} is higher than other values, or if microbial C is higher than normal range, and DT_{50} is lower than other value, reliability is questionable. if microbial C changes during test by factor of 2, DT_{50} should be evaluated	E
2. DT_{50} of metabolites	if for metabolites > 10 % no reliable DT_{50} can be calculated, additional transformation studies are required	

Notes

a If the application rate is more than two times too high: inhibition (or stimulation) of the soil organisms is possible, leading to a higher (lower) DT_{50}; if the application rate is more than two times too low: inhibition of the soil organisms might occur at the normal application rate.

b See methods of Van Vlaardingen and Smit (2008) for the estimation of pF. Soils should not become too wet or too dry (pF should be in the range 2 - 4), also during pre-treatment. In case of deviating pF, DT_{50} should be judged in relation to other available data.

5.3 Sorption and mobility in unsaturated soil

5.3.1 General aspects

Sorption may be studied through experiments with:

- soil-water suspensions (also called adsorption, batch-equilibrium or slurry experiments);
- soil columns (also called leaching experiments);
- thin or thick layer chromatography (TLC) experiments.

In the respective sections below further instructions are given. Summary Table 5-2 refers to batch equilibrium experiments according to OECD 106.

If a pesticide has one or more acid or basic groups, the soil pH influences its mobility. As in Dutch soils the negative charges of soil particles prevail, it is important to verify whether a pesticide is dissociated (i.e. negatively charged, due to the loss of H^+), protonated (i.e. positively charged, due to the take up of H^+), or neutral, given a specific soil pH. The soil pH determines the extent of dissociation or protonation of such pesticides, and therefore their mobility. The potency of a pesticide to protonate or dissociate is expressed by its pK_a. Generally, negatively charged pesticides in a rather basic soil are assumed to be more mobile than neutral or protonated pesticides.

The extent of sorption is indicated by a distribution constant K_d, which is the ratio between the content of the substance in the soil phase and the mass concentration of the substance in the aqueous solution. In soil column leaching experiments the distribution constant is also indicated as $K_{s/l}$, because it concerns the distribution between the solid phase and the liquid phase. In batch equilibrium studies, the distribution constant is indicated as K_F as it derived from the Freundlich equation.

The K_{OM} value is derived from the $K_{s/l}$ value according to Equation 5.

$$K_{OM} = \frac{100 \times K_{s/l}}{\%OM} \qquad \text{Equation 5}$$

in which:

K_{OM} = organic matter partition coefficient [L/kg]
$K_{s/l}$ = solid-liquid partition coefficient [L/kg]
%OM = percentage organic matter of the soil

$K_{s/l}$ or K_{OM} values have to be checked or recalculated from the raw data, if available. The value calculated by the author is mentioned in Results, the recalculated values in the Remarks.

For matters of convenience, we accept the K_F as a $K_{s/l}$ for further calculations, provided that the Freundlich exponent is within the range 0.7 - 1.1. In case the 1/n exponent is outside the range 0.7 - 1.1, the applicant should be requested to submit additional information to explain the deviation and to give a justification of using the results for K_{OM} calculations.

In case transformation was too high in the adsorption experiment, a reliable $K_{s/l}$ can only be calculated if both the concentration in the liquid phase and the amount of the substance adsorbed to the soil are determined. When adsorption is solely based on the concentration decrease in the aqueous phase, accurate determination of the $K_{s/l}$ is only possible if $K_{s/l}$ × (ratio soil/water) > 0.3. When both soil and solution are analysed, an accurate $K_{s/l}$ can be derived when $K_{s/l}$ × (ratio soil/water) > 0.1.

Table 5-2 Soil, sorption and mobility, adsorption studies

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance and position of label	improperly characterised or reported?	Y [→Ri 3]
2. vehicle	reported? amount used?	E
3. soil[a]		Y
3.1 soil type	USDA-class and other relevant data reported? soil must be relevant for the agricultural situation (e.g. no paddy or volcanic soil; 0.5 – 15 % OM)	Y
3.2 characteristics	pH, CEC, OM reported?	Y
4. weight of soil sample	not reported?	E
5. soil/water ratio (g/mL)	not reported?	Y
6. temperature	not reported?	E
7. number of concentrations	min. X concentrations [X = 5] should be used	Y
8 number of replicas	test should be performed in duplo	Y
9. shaking time	shaking time (in hours) should be shorter than the DT_{50} (in days)[b]; no longer than 48 hours	E
10. extraction/analysis method	should be appropriate for the substance and the metabolites, and the recovery of the substance should be >X% [X = 70] and <Y% [Y = 110]	E
RESULTS		
1. mass balance	recovery should be > X% [X = 80]	Y
2. distribution constants	$K_{s/l}$, K_F, K_{OM}, K_{OC} $K_{s/l}$ is only accurate if $K_{s/l}$ x (soil/water ratio) > 0.3 (water fraction analysed only) or > 0.1 (water and soil analysed)	E
3. Freundlich exponent	K_F with 1/n outside range 0.7 – 1.1 should be verified	Y
PAY ATTENTION TO		
1. water solubility	initial and equilibrium concentrations should not exceed water solubility	Y
2. transformation	there should be no major loss due to transformation (max. 3%), unless both the amount sorbed and the decline in concentration of the substance in the liquid phase has been determined	E
3. pK_a of the substance	if 2 < pK_a < 6: $K_{s/l}$ should preferably be determined at pH 7-8, or in soils with pH at least 2 units higher than pK_a	Y
4. soil handling	no manipulation, with exception of sieving (2 mm) allowed. Sterilisation is not allowed	Y

Notes
a For some pesticides (e.g. those containing a phosphate moiety) the amounts of sesqui-oxides/hydroxides in soil might explain the amounts sorbed. For some pesticides (e.g. those with a positive charge) the amounts of clay might explain the amounts sorbed.
b This requirement has to be met if only the decline in concentration of the substance in water is analysed and not the amount sorbed to the soil: if the shaking time (in hours) is longer than the DT_{50} (in days), the concentrations in both soil and water must be measured.

5.3.2 Column studies with unaged test substance (soil)

Summary Table 5-3 refers to column leaching studies. In these experiments the penetration depth of the substance is generally determined. This depth is estimated as the depth of the bottom of the layer above which half of the substance recovered (in soil column plus leachate) is present. With the penetration depth the sorption coefficient can be calculated. The penetration depth (X_p) is established as follows:

- The substance is applied on top of the soil column without homogeneous mixing:
    - If 50 % or more of the applied compound is found in the column X_p is established as the distance from the top of the column to the bottom of the layer in which the 50% is reached.
    - If more than 50 % is in the leachate and the leachate is not analysed in fractions, X_p is defined as ≥ column length; consequently the K_{om} cannot be calculated accurately and a ≤ sign is added (and the reliability is lowered).

- If more than 50 % is found in the leachate and the leachate is analysed in fractions, X_p equals the column length; the thickness of the water layer (D) used in the formula below is adapted. In this case we do not need the thickness of the totally applied water layer, but only the thickness of the water layer that leached 50% of the substance. The thickness of this water layer is calculated from the amount of leachate that contains the first 50% of the substance. The volume of this leachate is divided by the column area to give the thickness of the water layer.

- The substance is homogeneously mixed with soil and applied in a soil layer on top of the soil column (this applies also to soil with aged residue):
  - If 50 % or more is found in the applied soil layer, the X_p is established as half the height of the applied soil layer, resulting in a $K_{s/l}$ value with a '≥' sign.
  - If 50 % is found in a layer below the applied layer, the penetration depth is established as the distance to the bottom of the segment in which the 50% limit of the penetrated substance is reached minus half the thickness of the applied soil layer.
  - If more than 50 % is found in the leachate and the leachate is analysed in fractions, X_p equals the column length minus half the thickness of the applied soil layer; the thickness of the water layer (D) used in the formula below is adapted in the same way is described above).

The penetration depth X_p and the $K_{s/l}$ are related as follows (Equation 6 and 7):

$$X_p = \frac{D}{\left(\rho \times K_{s/l}\right) + \Theta}$$ Equation 6

and

$$K_{s/l} = \frac{\frac{D}{X_p} - \Theta}{\rho}$$ Equation 7

in which:
X_p = the penetration depth of the substance [cm]
D = the thickness of the applied water layer [cm]
ρ = the dry soil bulk density (default 1.5) [kg/dm^3]
$K_{s/l}$ = the solid-liquid partition coefficient [dm^3/kg]
θ = the moisture volume fraction (default 0.4) [-]

If the water movement is followed using a tracer (Equation 8):

$$X_p = \frac{\theta \times X_{p,tracer}}{(\rho \times K_{s/l}) + \theta}$$ Equation 8

in which:
X_p = the penetration depth of the substance [cm]
$X_{p,tracer}$ = the penetration depth of the tracer [cm]
ρ = the dry soil bulk density (default 1.5) [kg/L]
$K_{s/l}$ = the solid-liquid partition coefficient [L/kg]
θ = the moisture volume fraction (default 0.4) [-]

In BBA-37 studies (BBA, 1986), the leaching time is two days. In case in BBA-37 studies the percentages organic matter (OM) are not reported, the following values are used for the soils 2.1 (sand), 2.2 (loamy sand), 2.3 (sandy loam): 1%, 4%, and 2%, respectively. In general, BBA-37 studies get the Ri 2, because the columns are not sliced.

Table 5-3 Soil, sorption and mobility, column leaching studies

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance + position of label	improperly characterised or reported?	Y [→Ri 3]
2. vehicle	reported? amount used?	E
3. soil[a]		Y
3.1 soil type	USDA-class and other relevant data reported? soil must be relevant for the agricultural situation (e.g. no paddy or volcanic soil; 0.5 – 15 % OM)	Y
3.2 characteristics	pH, CEC, OM reported?	Y
4. aged or unaged treated soil	aged soil: amount of a.i. and metabolites present after ageing should be determined	Y
4.1. ageing time	give duration, in general equal to the DT_{50} of the most relevant compound	E
5. application	way of application reported?	Y
5.1 rate	not reported?	Y
5.2 way of application	unaged: e.g. spraying aged: it is important that the applied amount of aged soil is homogeneously mixed	E
6. soil column		
6.1 length	length min. X cm [X = 20]	E
6.2 diameter	∅ min. X mm [X = 25]	E
7. water saturation prior to leaching	column should be at field capacity before the test substance is applied	E
10. waterflux	maximum X [X = 10] cm/day	Y
11. leaching time		
12. samples[b]		
12.1 column slices	column should be sliced: min. X layers [X = 5], with max. thickness of X cm [X = 5 cm]	Y
12.2 leachate	should be analysed in at least X fractions [X = 5]	Y
13. extraction/analysis method	should be appropriate for the substance and the metabolites, and the recovery of the substance should be >X% [X = 70] and <Y% [Y = 110]	E
RESULTS		
1. penetration depth		
2. total recovery (maximum, and time)	specify: of applied or after ageing. Recovery of r.a. after ageing of the activity should be >X% [X = 80]	E
3. distribution of active and metabolites over the column/leachate	give percentages, preferably in a table	
4. metabolites	metabolites identified and quantified separately? give identity and percentages	Y
PAY ATTENTION TO		
1. pK_a of the substance	if 2 < pK_a < 6: sorption should preferably be determined at pH 7-8, or in soils with pH at least 2 units higher than pK_a	
2. handling of soil	no manipulation with exception of sieving (2 mm) is allowed. According to the BBA IV 4-2 Guidelines the sieve diameter is 1 mm	E
3. volatilisation	there should be no major loss due to volatilisation	E

Notes

a For some pesticides (e.g. those containing a phosphate moiety) the amounts of sesqui-oxides/hydroxides in soil might explain the amounts sorbed. For some pesticides (e.g. those with a positive charge) the amounts of clay might explain the amounts sorbed.

b In general both column and leachate should be analysed in layers or fractions, but:
- if < 25% in leachate: the column should be sliced, but the leachate need not be analysed in fractions;
- if > 75% in leachate: the column need not be sliced, but the leachate should be analysed in fractions;
- in all other cases: both column and leachate should be analysed in fractions.

5.3.3 Column studies with aged test substance (soil)

In general, studies with freshly applied residues are used. Aged-residue tests are generally performed to obtain qualitative information on the behaviour of metabolites. These tests are also included in Summary Table 5-3. In some cases studies with aged residue may be used to obtain a K_{OM}, if the penetration depth of the substance can be determined in a reliable way. After ageing, the soil must be analysed for individual substances. Instructions for determining $K_{s/l}$ values from aged-residue studies (for actives and metabolites) are under development.

5.3.4 TLC studies (soil)

Soil TLC stands for both soil thin layer chromatography and soil thick layer chromatography. In these studies, the retardation factor Rf is determined, which is the ratio between the elution distance of the substance and the elution distance of the developing phase. The Rf is used to derive the solid-liquid partition coefficient for the substance: $K_{s/l}$. Soil TLC studies are not frequently submitted anymore. $K_{s/l}$ values determined in TLC experiments are considered less reliable because of difficulties in the exact determination of the relative rates of movement. The sorption may be underestimated, due to handling of the soil, possible influence of the support material, and a probable non-equilibrium situation. Therefore, the highest Ri these studies can get is 2, and the result is not used for risk assessment.

5.3.5 Field tests (lysimeter studies)

Detailed guidance for summarising and evaluating lysimeter studies is given in Verschoor et al. (2001).

5.4 Fate and behaviour in surface water

5.4.1 General aspects

If raw data are available, the DT_{50} values always have to be checked or recalculated. At least five time points including the value on t = 0 are considered necessary. The DT_{50} value can generally be calculated by non-linear regression of first-order kinetics. If $r^2 < 0.8$, the regression is not considered valid. If model fitting is not possible, the results are reported as DT_{50} > x days or < y days, depending on the data.
In the Remarks should always be stated, if the transformation followed first-order kinetics or not, and if the DT_{50} was calculated or estimated otherwise. In case the DT_{50} was extrapolated, this should also be mentioned in the Remarks.

All hydrolysis, photolysis, or biodegradation (i.e. respecting the whole water/sediment system) DT_{50} values are converted to 20 °C using Equation 4.

5.4.2 Photolysis studies (surface water)

The instructions in Summary Table 5-4 are based on the EPA-540/9-85-014 Guideline.

Table 5-4 Water, phototransformation

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance + position of label	improperly characterised or reported?	Y [→Ri 3]
2. applied concentration(s)	not reported? test concentrations should not exceed water solubility	Y
3. use of vehicle	reported? amount used?	E
4. buffer system	buffer solution has to be used unless it has been proven by measurement that pH is not changing	Y
5. pH	checked and maintained?	E
6. sterility	checked and maintained?	Y
7. use of sensitiser	reported? amount used?	E
8. oxygen	excluded: yes or no	E
9. radical scavenger	used: yes or no	E
10. test duration		
11. temperature	monitored? constant?	Y
12. type of light; light intensity; light distance	artificial light should be comparable with sunlight (> 290 nm)	Y
12.1 dark control	not included?	Y [→Ri 3]
13. sampling frequency	sampling should be frequent [≥ 5] enough to establish DT_{50} and kinetics	E
14. extraction/analysis method	should be appropriate for the substance and the metabolites, and the recovery of the substance should be > X% [X = 70] and < Y% [Y = 110]	E
RESULTS		
1. total recovery (maximum, and time)	recovery should be >X% [X= 80]	Y
2. amount of parent and metabolites over time	give table or numbers	
3. metabolites	metabolites identified and quantified separately? give identity and percentages	Y
4. dark control	substantial degradation?	E
5. quantum yield	if reported	
6. DT_{50}	give way of calculation in case of discontinuous exposure: it should be stated, if DT_{50} is calculated over the whole period, or only over sun hours	
PAY ATTENTION TO		
1. the dissipation type	there should be no major loss due to biotransformation (sterile conditions should be maintained), photolysis, volatilisation, or adsorption to test vessel: is the test design adequate?	E
2. quantum yield	quantum yield Φ is calculated as the amount of substance (in mol) transformed per amount of photons (in Einstein) and has no dimension.	E
3. optical density	if the optical density > 1, the amount of light is the limiting factor in the transformation. Then the actual DT_{50} may be smaller.	E

5.4.3 Hydrolysis studies

The hydrolysis test (Summary Table 5-5) should be carried out at three different pHs: 4, 7, and 9. The hydrolysis at each pH has to be tested at 2 - 3 different temperatures, unless the substance is thermostable (see OECD 111). The hydrolysis rate can be influenced substantially by a small change in temperature, and the extent is dependent on the compound.

Table 5-5 Water, hydrolysis

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance + position of label	improperly characterised or reported?	Y [→Ri 3]
2. applied concentration(s)	not reported? test concentrations should not exceed water solubility	Y
3. use of vehicle	reported? amount used?	E
4. buffer system	buffer solution has to be used unless it has been proven by measurement that pH is not changing	Y
5. pH	preferably 4, 7, and 9	E
6. sterility	checked and maintained?	Y
7. test duration		
8. temperature	monitored? constant?	Y
9. light condition	test should be performed in darkness, unless it has been proven that photodegradation is of no importance in the same buffer solutions	Y
10. sampling frequency	sampling should be frequent [≥ 5] enough to establish DT_{50} and kinetics	E
11. extraction/analysis method	should be appropriate for the substance and the metabolites, and the recovery of the substance should be >X% [X = 70] and <Y% [Y = 110]	E
RESULTS		
1. total recovery (maximum, and time)	recovery should be > X% [X = 80]	Y
2. amount of parent and metabolites over time	give percentages, preferably in a table	
4. metabolties	metabolites identified and quantified separately? give identity and percentages	Y
3. DT_{50}	give way of calculation; indicate whether DT_{50} is extrapolated beyond test duration	
PAY ATTENTION TO		
1. the dissipation type	there should be no major loss due to biotransformation (sterile conditions should be maintained), photolysis, volatilisation, or adsorption to test vessel: is the test design adequate?	E

5.4.4 Water/sediment studies (surface water)

Water/sediment studies (Summary Table 5-6) are most often performed according to OECD 308 or BBA IV, 5-1. The water-sediment system has to acclimate in order to recuperate from the disturbance and to allow the formation of gradient zones, especially from aerobic to anaerobic.
For water/sediment systems, a distinction is made between the DT_{50} value for the pesticide in the aqueous phase ($DT_{50,water}$), the DT_{50} value in the sediment phase ($DT_{50,sediment}$), and the DT_{50} value for the whole water/sediment system ($DT_{50,system}$). If there are indications that dissipation from the water phase is determined by other processes than transformation (e.g. sorption to sediment), this should be indicated in the Header and in the Remarks (see also 5.2.1).

Results of sterile incubations are not presented in the Header, but are given in the Results. Unusual results (e.g. degradation) are mentioned in the evaluation.
If raw data are available, the DT_{50} values always have to be checked or recalculated. Detailed guidance on the (re)calculation of DT_{50} values is given in FOCUS (2006).

Table 5-6 Water, transformation in water/sediment systems

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance + position of label	improperly characterised or reported?	Y [→Ri 3]
2. applied concentration(s)	not reported? test concentrations should not exceed water solubility	Y
3. use of vehicle	reported? amount used?	E
4. test system		Y
4.1 origin and type of sediment	USDA-class and other relevant data reported?	Y
4.2 characteristics	pH, OM[a], redox potential reported?	Y
4.3 acclimatisation	at least 6 - 8 weeks	E
4.4 amount of water and sediment	the mass fraction of solids (dry weight) in the system should be X% based on total weight system [X= 1 or 10[b]] an X cm sediment layer should be used: [X= 2] for small shallow water (agricultural pesticide) [X= 0.2] for large water (non-agricultural pesticide)	Y
4.5 redox potential	monitored? water should be in aerobic condition[c], sediment should be anaerobic	Y
4.6 volatile traps	present?	
5. temperature	monitored? constant?	Y
7. light conditions	not reported?	Y
8. test duration	until 90% of a.i. is transformed, up to 100 days	Y
9. sampling frequency	sampling should be frequent [≥ 5] enough to establish DT_{50} and kinetics	Y
10. extraction/analysis method	should be appropriate for the substance and the metabolites, and the recovery of the substance should be > X% [X = 70] and < Y% [Y = 110]	Y
RESULTS		
1. recovery and distribution of radioactivity over time	total recovery should be > X% [X = 80]	Y
1.1 radioactivity: total, extractable, bound residues (BR), CO_2, volatiles	present table	
1.2 radioactivity assigned to parent and metabolites	identified and quantified separately? present table	Y
1.3 relevant metabolites	give name and maximum; relevant: > 10 % of AR, > 5 % of AR at 2 consecutive endpoints or maximum at end	
2. $DT_{50,water}$, $DT_{50,sediment}$, $DT_{50,system}$	give way of calculation indicate whether $DT_{50,water}$ is dissipation or degradation	
PAY ATTENTION TO		
1. dissociation of the compound	the pH may influence the partitioning between water and sediment	E

Notes
a Two types of sediment have to be tested for agricultural pesticides: one with a high, and one with a low % OM. For substances that will mainly reach larger surface waters (e.g. non-agricultural pesticides) only one system has to be tested (< 5 % OM).
b A higher percentage of sediment might reduce the biotransformation as there could be a high degree of sorption.
c If the water is aerated, the sediment should not be disturbed.

5.4.5 (Ready) biodegradability

The results of a Readily Biodegradability test are used to estimate the rate constants for elimination of a compound in a Sewage Treatment Plant (STP). Details for testing and interpretation of results are given in OECD 301. The criteria for the classification of a substance as 'readily biodegradable' are given in the Technical Guidance Document (TGD; EC, 2003).

For two additional tests, the following criteria apply for a substance to be classified as ‘readily biodegradable’ (J. Struijs, RIVM, pers. comm.):

ISO 7827:	> 70% DOC removal
Simulation (OECD 303A, 1981):	> 80% DOC removal

If a substance is not degradable and the toxicity control shows < 25 % degradation (on the basis of ThOD of $ThCO_2$ of < 35% degradation on the basis of total DOC after 14 days, the substance is inhibiting bacteria and testing the effects on bacteria in activated sludge is required (OECD 224). If a substance is toxic for micro-organisms, this should be accounted for in the ‘Ready test’.

5.4.6 Field tests

In some cases, aquatic model ecosystems (micro- and mesocosms) are submitted for registration purposes. These tests are most often performed to obtain additional information on the aquatic ecotoxicity, but can also be used to evaluate the fate of a compound under (semi-)field conditions. The summary and evaluation of the field test should support the objective of risk assessment. Guidance on performance of mesocosm studies is given in a series of publications, e.g. OECD (2006b); Campbell et al. (1999); Hill et al. (1994); Arnold et al. (1991).

6 Ecotoxicity

6.1 General aspects

In case formulations are tested, the result is presented in the Header on the basis of the product and on the basis of the active ingredient(s).

6.1.1 Statistics

Detailed guidance on the statistics to be applied for ecotoxicity testing is published by OECD (OECD, 2006a). For the calculation of the $L(E)C_{50}$ or LD_{50} values, commonly used methods such as probit, moving average angle and linear interpolation are accepted, but the reported values should be checked with the raw data. If questions arise about the reported value, or in case actual concentrations must be used instead of nominal, endpoints have to be recalculated. The $L(E)C_{50}$ or LD_{50} values for discrete variables (mortality, immobilisation) can be estimated according to Spearman-Kärber (Hamilton et al., 1977). For continuous variates, such as (algae) growth, non-linear regression of a sigmoid concentration-response curve can be applied (see e.g. Haanstra et al., 1985; Van Ewijk and Hoekstra, 1993) In all cases, it should be mentioned in the Remarks whether endpoints were recalculated and which method was used.

6.2 Birds

6.2.1 Acute oral toxicity (LD_{50})

There is no OECD guideline for acute studies, most studies are performed according to OPPTS guidelines. For the oral acute LD_{50}-studies, test duration is not entered in the Header, because the test substance is administered in a single dose.

In acute studies, birds may vomit due to the high dose that is applied, resulting in an underestimation of the ingested dose. If the birds vomit, this should be reported under Results. In case vomiting occurred, but no mortalities are observed, the LD_{50} is assumed to be higher than the highest concentration at which *no* vomiting occurred. The concentration with vomiting are not used for (re)calculation of the LD_{50}.

6.2.2 Short-term toxicity (LC_{50})

In case of an eight-days sub-acute study (OECD 205), the exposure time is five days and the observation period three days.

It should be checked whether consumption in the test item groups is similar to that of the control, and findings should be reported. In case of decreased consumption related to concentration, the No Repellent Concentration (NoRC) should be estimated according to Luttik (1993).

The LC_{50} should be reported on the basis of concentrations in feed (mg/kg fd) and on the basis of the daily dose (mg/kg bw.d). The general rule for the conversion is:

$$Dose = \frac{C_{food} \times Cons}{BW}$$ Equation 9

in which:

Dose	=	daily dose [mg/kg bw·d]
C_{food}	=	concentration in food [mg/kg fd]
Cons	=	daily food consumption [g/bird·d]
BW	=	body weight [g]

According to the Guidance Document on risk assessment for birds and mammals under Directive 91/414/EEC (EC, 2002a), the following data should be used for this conversion:

- For food consumption, usually group consumption rates (expressed as g per bird per day) are given in the report for the 5-day-exposure period and the 3-day-post exposure period; the former figure is needed here.
- For body weight, the group means for day 0, 5, and 8 are reported. For the purpose here take the average of day 0 and day 5.

The conversion from concentration to daily dose is not appropriate for those treatment groups where a strong food avoidance is obvious (in that case the average dose over 5 days is misleading) as well as for treatment groups with a high mortality (in that case data for the body weight at day 5 and for the food consumption have a poor quality or are missing at all). Three cases are distinguished:

- Case 1: LC_{50} is above top concentration
  Convert each treatment group separately (however, only the top level is needed for the risk assessment).
- Case 2: LC_{50} is below top concentration, food consumption not affected.
  There are two possibilities:
  - convert concentration into achieved dose for each treatment group, and conduct a new probit analysis, this time using the daily-dose data
  - take the overall mean value for food consumption and body weight (mean from all dose groups where calculation is possible) and use these figures to convert the LC_{50} (this option is sensitive against concentration-dependent food avoidance).

  If food consumption is slightly affected expert judgement is required to decide whether procedures according to case 3 should better be followed.
- Case 3: LC_{50} below top concentration, distinct food avoidance well below the LC_{50}.
  Conversion from concentration to daily dose may be unreliable due to the low number of survivors on which food consumption is based; furthermore food consumption may change markedly from day to day. These problems alone should be no reason to repeat the study. Rather assessment should be conducted on a case-by-case basis (e.g. if the study delivers a NOEC then this could be a starting point as the converted LC_{50} must be well above this level); expert judgement is needed in this case to decide on the endpoint to be taken for risk assessment.

Table 6-1 Birds, acute and subacute toxicity studies

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance	improperly characterised or reported?	Y [→Ri 3]
2. test species	not reported?	Y [→Ri 3]
2.1 sex, weight and age of the birds	not reported?	Y
2.2 acclimatisation, starvation period	not reported?	Y
3. application	not reported?	Y [→Ri 3]
3.1 type of application	oral (by e.g. gavage), or dietary; not reported?	Y
3.2. concentrations	min. X concentrations [X = 5]; limit test allowed if LD_{50} > 2000 mg/kg bw or LC_{50} > 5000 mg/kg fd (range finding)	E
3.3. use of vehicle	reported? amount used?	E
4. number of animals per group	not reported?	Y
5. feed type (LC_{50} study)	not reported?	Y
6. exposure time (LC_{50} study)	preferably X days [X = 5]	E
7. observation time	LD_{50}: c. X days [X = 14] LC_{50}: after five days exposure X days of observation [X= 3]	E
8. availability of feed and water	should be available *ad libitum*	E
9. housing conditions	group or individual	
RESULTS		
1. LC_{50} or LD_{50}, and 95% confidence limits	raw data available? LC_{50} in [mg/kg fd] and as daily dose [mg/kg bw.d]	Y
2. mortality data in all groups	the mortality in the controls should not exceed X% [X= 10] at the end of a test	Y
3. sublethal effects (overt signs of toxicity and macroscopic effects)		
4. feed consumption	not reported? indication of repellency (see 8); consumption needed to recalculate daily dose in LC_{50} test	Y
5. body weight change	not reported? body weight day 1 – 5 needed to recalculate daily dose in LC_{50} test	Y
6. measured concentrations (LC_{50} study)	not measured? actual test concentrations should > X % of nominal [X=80]	E
7. vomiting	if vomiting occurs the actual toxicant intake is unknown; LD_{50} is lowered	E
8. repellency	if repellency is claimed, or if feed consumption decreases with increasing test concentrations, a NoRC may be estimated, see Luttik (1993)	E
PAY ATTENTION TO		
1. LC_{50}: stability/homogeneity of the substance in the diet	stability/homogeneity should be maintained throughout the test	Y
2. LC_{50}: mortality at the lowest concentration	at the lowest concentration no toxic effects should appear	E

6.2.3 Reproductive toxicity (NOEC)

In chronic tests (according to OECD 206) the exposure time is at least 20 weeks, and the following effects are studied: mortality of the adults, egg production, cracked eggs, eggshell thickness, viability, hatchability and effects on young birds. All parameters can be used for the derivation of the NOEC, but a NOEC cannot be based on eggshell thickness if this is the only parameter showing an effect.
It should be checked whether there are apparent non-breeders in the control group. Data from these replicates should be left out of the statistical analyses.

For the conversion of the NOEC from mg/kg fd to a daily dose (see Equation 9), each treatment group is converted separately. The following input should be used:

- Feed consumption: Data are reported on a weekly basis for pairs or groups. Feed consumption usually is higher during egg-laying (to be attributed to the females), however, for the purpose here the average consumption over the entire exposure period is taken. Separate calculations are performed for males and females.
- Body weight, the average body weight for both sexes over the exposure period is taken.

6.3 Aquatic organisms

6.3.1 Nomenclature

For species of which the scientific (Latin) name has been changed, the current name is used. This may apply e.g. to rainbow trout (*Oncorhynchus mykiss*, formerly *Salmo gairdnerii*) and the green algae *Selenastrum capricornutum* (now *Pseudokirchneriella subcapitata*).
Blue algae are prokaryotes and are classified as Cyanophyta, as to distinguish them from the eukaryotic green algae, which are classified as Chlorophyta.

6.3.2 Maintenance and analysis of test concentrations

There should be no major loss due to degradation, photolysis, volatilisation, hydrolysis, adsorption to glass, etc. The test design should be adequate as to maintain >80% of the nominal concentration, where needed and if possible, a renewal or flow-through system should be used (see also 6.3.3). Although for acute tests chemical analysis of test water is not obligatory according to the guidelines, verification of test concentration is strongly recommended. In case photolysis occurs, tests with *Daphnia* and fish may be performed in the dark.

For tests with slightly soluble substances or rapidly hydrolysing substances some additional instructions are mentioned below.

6.3.2.1 Tests with slightly soluble substances

In case very slightly soluble compounds (water solubility S <0.1 mg/L) are tested at concentrations up to the water solubility, and no effects are observed, the test is in principle reliable. Toxicity values from a test in which the slightly soluble compound was tested at nominal concentrations that are larger than 10 times the solubility, should be regarded as unreliable (Ri 3). The reader is referred to Vaal et al. (1992) for more information on the evaluation of slightly soluble substances.

6.3.2.2 Tests with rapidly hydrolysing pesticides

The EC and OECD Test Guidelines have been devised for stable compounds. However, one can be confronted with an unstable pesticide. According to the official Guidelines, the loss of test substance in an ecotoxicological test should be less than 20 % to consider the compound stable enough for the purposes of toxicity testing. If the loss of test substance is higher than 20 %, first it should be checked if the loss is caused by bad performance of the test (in which case the Ri is lowered), or by fast hydrolysis. The latter is done by inspecting the results of the hydrolysis test, and if these are not available (or in case of doubt), by consulting a specialist. Since high hydrolysis rate is an intrinsic property of the compound, it cannot be avoided. If all possible measures have been taken to maintain test concentrations (e.g. renewal; flow-through), and the test was thus performed in a technically adequate way, the test is considered reliable. A > 20 % loss may thus be acceptable in some cases.

Once it is established that hydrolysis is the (main) cause for the high loss, the second question is if the metabolite(s) should have been tested instead of the parent compound (of course it can also be concluded that testing of both parent compound and metabolite(s) is necessary). The following limits are used as a rule of thumb[2]:

$DT_{50,hydrolysis} \geq$ 24 h : the test is started with the parent substance
$DT_{50,hydrolysis} <$ 4 h : test(s) are started with metabolite(s)
$DT_{50,hydrolysis}$ 4-24 h : expert judgement

In case the $DT_{50,hydrolysis} <$ 4 h, and the toxicity test has been started with the parent compound, the test is considered not suitable to assess the toxicity of the parent, but may be considered valid to meet the data requirement for metabolite testing.

In all other cases (i.e. loss of test substance > 20 %, metabolites neither qualified nor quantified, $DT_{50,hydrolysis} \geq$ 4 h), the test results are considered as less reliable (Ri 2), as it remains unclear which compound causes the observed effect(s) and at what concentration. The toxicity is then expressed in terms of the nominal (initial) concentration[3], and a comment is added in the Remarks that the toxicity is determined by a mixture of the parent compound and one or more transformation products, because of rapid hydrolysis of the parent compound. The usefulness of such test results for risk assessment depends on the situation to be assessed.

6.3.3 Calculation of test endpoints: nominal or measured concentrations

In the Guidance Document on Aquatic Ecotoxicology (EC, 2002b), the following guidance is given with respect to the use of nominal or measured concentration for the calculation of endpoints:
Toxicity endpoints (LC/EC_{50}, NOEC, etc.) should usually be calculated using nominal concentrations, because these are most suitable for comparison with maximum Predicted Environmental Concentrations (PECmax). In studies where the initial measured concentrations are < 80 % or > 120% of the nominal, toxicity values should be presented as measured initial concentrations. These can then be used in the risk assessment in the same way as nominal concentrations. This approach is especially relevant for static tests.
If the measured concentrations are < 80 % or > 120 % of nominal ones during the test, toxicity values should be presented additionally as measured concentrations and the mean measured concentration for the relevant test period should be used to express toxicity. If the measured concentrations are very low compared to nominal ones, the validity of the test might be questionable, and a justification for using such study should be required. In summary, the following rules apply in general:

- if measured concentrations are > 80 %, then the nominal concentration can be used to express toxicity,
- if initial measured concentrations are < 80 %, then toxicity values should be expressed as initial measured concentrations,
- if measured concentrations in semi-static and flow-through systems fall graduately below 80 % during the test, then toxicity values should be expressed as mean measured concentrations.

[2] The limits are adapted with modifications from Whitehouse and Mallet (1993).
[3] Whitehouse and Mallet (1993) use the term ‘loading rate’.

6.3.4 Toxicity to algae, *Daphnia* and fish

Summary Table 6-2 refers to acute tests with algae (OECD 201), *Daphnia* (OECD 202) and fish (OECD 203). For algae, it should be noted that the former 1984-version of OECD 201 included both growth rate and biomass, calculated as the area under the growth curve, as endpoints. In the revision of 2006, growth rate is preferred as endpoint, yield (increase in cell numbers) may be calculated as an additional endpoint. It should also be noted that in most study reports both an EC_{50} and NOEC are derived. For risk assessments according to the TGD (EC, 2003), the EC_{50} is considered as an acute endpoint, and the NOEC as chronic. In the context of 91/414/EEC, only the EC_{50} was used as the input for risk assessment, but this may be changed into using a NOEC upon revision of the Directive.

Most items in Table 6-2 are applicable to chronic tests as well, although there are important differences between short-term and long-term tests. Naturally, the exposure time is longer in the latter: 21 days for *Daphnia*, and 28 days for a fish juvenile growth test, or covering the life cycle from eggs to free feeding juveniles in an Early Life Stage test (ELS). Further, the test system has to be either semi-static or flow through, never static. Also, feeding is allowed/required, in contrast with short-term tests. The studied effects are:

Daphnia: effects on mortality, time of first production of young, number of young per surviving female, signs of intoxication;

fish (growth): growth of juvenile fish;

fish (ELS): effects on the stage of embryonic development, hatching and survival, abnormal appearance, abnormal behaviour, weight, and length (darkened skin is not considered relevant).

For more detailed information, the reader is referred to OECD 211 (*Daphnia*, reproduction), OECD 210 (fish, Early-Life Stage) and OECD 215 (juvenile growth).

Male daphnids may indicate bad culture conditions, and influence the number of offspring as well as the statistical analysis. Check the effect of the number of males on the calculated NOEC, as the latter should be based on females.

Table 6-2 Water organisms, short-term toxicity tests

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance	improperly characterised or reported?	Y [→Ri 3]
2. test species	not reported?	Y [→Ri 3]
2.1 age (crust.+fish); length+weight (fish)	not reported? *Daphnia*: maximum age 24 h	Y
2.2 acclimatisation	not reported?	Y
3. test system	static, renewal, flow through not reported?	 Y
4. test water/medium	not reported?	Y
4.1 source	good quality natural water or reconstituted water	E
4.1 characteristics	hardness 10-250 mg $CaCO_3$/L; pH 6.0-8.5	E
5. number of animals/ loading		
5.1 algae	algae: initial cell density 1 x 10^4 cells/mL	E
5.2 *Daphnia*	min. 20/conc.: preferably 4 replicas of 5; max. 1 *Daphnia* per 2 mL	E
5.3 fish	min. 7/conc.; max. 1 g/L (in flow-through tests the loading can be higher)	E
6. application		
6.1 concentrations	≥ X concentrations [X = 5] test concentrations should not exceed water solubility >X times [X = 10] limit test allowed if L(E)C_{50} > 100 mg/L or no mortalities at maximum water solubility	E
6.2 use of vehicle	concentration vehicle < 100 mg/L vehicle control should be included	E
7. exposure time	algae: 72-96h; *Daphnia*: 24-48h; fish: 96h	E
8. conditions		
8.1. light	not reported? algae: source, continuous 120 µE/m^2s ≥ 8000 lx; *Daphnia*: optional; fish: 12-16 h light/d	Y
8.2 temperature	see guideline; not reported?	Y
8.3 feeding	no feeding allowed in acute studies. Not reported?	Y
10. effects studied	algae: growth rate is preferred, yield is optional (OECD 201, revision 2006); *Daphnia*: immobility; fish: mortality and sublethal effects	E
11. analysis	not obligatory for acute tests, but strongly recommended	E
11.1 method	recovery > X % [X = 70] and < Y % [Y = 110]	E
11.2 sampling frequency	at least at the start and at the end of test	E
RESULTS		
1. measured concentrations	initial measured concentrations should be > X % [X = 80] and < Y % [Y = 120] of nominal	Y
2. pH, DO, temperature	pH and T should be constant; DO should be ⩾ 60% of air saturation value; see Appendix 1. report whether water quality parameters were within accepted range; list exceptions under Remarks	E
3. raw data	raw data should be available	Y
4. control mortality/effect	mortality/effect in the control should be <10% (or <1 fish if 7, 8, or 9 fish are used)	Y
5. endpoints	recalculate if necessary	E
5.1 algae NOEC and EC_{50} (95 % CI)	growth rate is preferred	
5.2 *Daphnia* EC_{50} (95 % CI)		
5.3 fish LC_{50} (95 % CI)		
PAY ATTENTION TO		
1. dissipation	there should be no major loss due to hydrolysis, photolysis, volatilisation, or adsorption to glass. Is the test design adequate?	E
2. log K_{ow}	bioconcentration, adsorption to glass and particle in solution may occur for lipophilic compounds	E
3. effects at lowest and highest test concentration	lowest: no toxic effect should appear; highest: algae: at least 50% inhibition; Daphnia: 100% immobilisation is preferred; fish: no percentage mortality is mentioned.	E

6.3.5 Bioconcentration in aquatic organisms

Summary Table 6-3 lists the items of interest for studies into bioconcentration in fish. The Bioconcentration Factor (BCF) should be based on the whole organism wet weight (BCF_{ww}) or on fat weight (BCF_{fw}). Studies based on OECD 305 A - E Guidelines should be checked with these Guidelines for evaluation. These Guidelines differ from each other with respect to the test system and the mathematical interpretation of the results.

Table 6-3 Bioconcentration in waterorganisms

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance (+ position of label)	improperly characterised or reported?	Y [→Ri 3]
2. test species 2.1 name, age, length, weight 2.1 loading	not reported?	Y [→Ri 3]
3. test system	static, semi-static or flow through	
4. test water/medium 4.1 source 4.1 characteristics	not reported? good quality natural water or reconstituted water hardness 10-250 mg $CaCO_3$/L; pH 6.0-8.5	Y E E
5. application 5.1 number of concentrations 5.2 range 5.3 vehicle	 min. X concentrations [X = 1] highest concentration < 0.1 x LC_{50}; lowest concentration > 10 x detection limit should not exceed 0.1 mL/L; vehicle control should be included	 Y E
6. exposure and depuration time	uptake phase: 3 hours - 30 days depuration phase: 6 hours - 60 days, or 3 x DT_{50}	E
7. conditions 7.1. light 7.2 temperature 7.3 feeding	 temperature should not vary more than ± 1 °C	 E
8. sampling frequency 8.1 water 8.2 animals	 uptake phase: min. X [X = 4] depuration phase: min. X [X = 5]	 E
9. analysis method	metabolites identified ?	Y
RESULTS		
1. signs of toxicity	no toxic effects should occur	Y
2. actual concentrations in water	should be ≥ 80% of nominal	
3. steady state concentration 3.1 concentration in water and fish	steady state reached: yes/no, time point concentration of active substance, not total radioactivity	 Y
4. BCF	based on whole body wet weight or lipid content	E
5. rate constants 5.1. k_2 and time to steady state	give k_1 (uptake) and k_2 (depuration); duration of phases should be sufficient time to reach 50 % of the steady state C_{fish} should be within a factor of 2 of $DT_{50,depuration}$	 E
PAY ATTENTION TO		
1. test concentrations	test concentrations should not exceed water solubility, and should be <1 mg/L	E
2. loss of test substance	there should be no major loss due to hydrolysis, photolysis, biotransformation in water, volatilisation, and adsorption to vessel or particles: is the test design adequate?	E

The OECD Guidelines 305 B-E use the model mentioned below. Guideline 305 A uses a different model, therefore the equations mentioned below cannot be used for 305 A. The Guidelines 305 B-D consider only the calculation of the BCF. When in these tests no steady state is reached, no BCF can be determined, which means that the result is not useful for risk assessment. This does not apply for the 305 E Guideline, because the rate constants that are determined in this test give insight in the behaviour of the chemical in the environment. The mass balance in the system consisting of water, organism (fish, or other) and test compound is:

$$\frac{dC_f}{dt} = (k_1 \times C_w) - (k_2 \times C_f) \qquad \text{Equation 10}$$

in which:
C_f = concentration of the test compound in the organism [µg/kg]
C_w = concentration of the test compound in water [µg/L]
t = time [d]
k_1 = uptake rate constant [L/kg·d]
k_2 = elimination rate constant [1/d]

The elimination rate constant k_2 describes every elimination process of the test compound from the organism, hence it includes physico-chemical elimination and biotransformation.

Integration of Equation 10 is only possible when C_w is constant. When C_w declines it must be checked if the authors included this in their calculations of the rate constants. If not, the rate constants can be recalculated using BIOFIT (Gobas and Zhang, 1992). BIOFIT can also be used if the C_w is constant.

Assuming C_w is constant, the next calculations should be performed. Consider Equation 10 during the phase of initial uptake, C_f will be then negligible:

$$\frac{dC_f}{dt} = k_1 \times C_w \quad or \quad C_f = k_1 \times C_w \times t) \qquad \text{Equation 11}$$

in which:
C_f = concentration of the test compound in the organism [µg/kg]
C_w = concentrationof the test compound in water [µg/L]
t = time [d]
k_1 = uptake rate constant [L/kg·d]

However, it is difficult to determine when C_f ceases to be negligible, so it is hard to use this formula in a good way. Another way to determine the uptake rate constant is described in OECD 305 E. For this method the k_2 is necessary:

$$\frac{dC_f}{dt} = (-k_2 \times C_f) \quad or \quad C_f = e^{(-k_2 \times t)} \qquad \text{Equation 12}$$

in which:
C_f = concentration of the test compound in the organism [µg/kg]
t = time [d]
k_2 = elimination rate constant [1/d]

When ln C_f is plotted against time, the slope of the straight line is k_2; k_2 can then be used to determine k_1. Integration of Equation 10 gives:

$$C_f = \frac{k_1}{k_2} \times C_w \times (1 - e^{(-k_2 \times t)}) \qquad \text{Equation 13}$$

in which:
C_f = concentration of the test compound in the organism [µg/kg]
C_w = concentration of the test compound in water [µg/L]
k_1 = uptake rate constant [L/kg·d]
k_2 = elimination rate constant [1/d]

k_1 is then calculated according to:

$$k_1 = C_f \times \frac{k_2}{C_w} \times (1 - e^{(-k_2 \times t)}) \qquad \text{Equation 14}$$

in which:
C_f = concentration of the test compound in the organism [µg/kg]
C_w = concentration of the test compound in water [µg/L]
t = time at which C_w and C_f are determined [d]
k_1 = uptake rate constant [L/kg·d]
k_2 = elimination rate constant [1/d]

Finally, for the bioconcentration factor, consider:

$$k_1 \times C_w = k_2 \times C_f \qquad \text{Equation 15}$$

in which:
C_f = concentration of the test compound in the organism [µg/kg]
C_w = concentration of the test compound in water [µg/L]
k_1 = uptake rate constant [L/kg·d]
k_2 = elimination rate constant [1/d]

From this equation it is clear that the bioconcentration factor (BCF in L/kg) is defined by:

$$BCF = \frac{k_1}{k_2} = \frac{C_f}{C_w} \qquad \text{Equation 16}$$

in which:
BCF = Bioconcentration Factor [L/kg]
C_f = concentration of the test compound in the organism [µg/kg]
C_w = concentration of the test compound in water [µg/L]
k_1 = uptake rate constant [L/kg·d]
k_2 = elimination rate constant [1/d]

The first part of this equation (BCF = k_1/k_2) can be used to determine the BCF also when no equilibrium in the test has been reached. The second part (BCF = C_f/C_w) can never be used to calculate BCF when equilibrium has not been reached.

The k_2 or half-life (DT_{50}) determines when equilibrium will be reached. The time required to reach 50 % of equilibrium concentration in the organism equals the half-life in the depuration part of the experiment. Whether the time to reach 50 % equilibrium and the half-life are comparable or not, should be checked in the original uptake curve. If this is not the case, the Ri of the test is lowered.

Equation 10 considers first-order, one-fish compartment kinetics. However, more-fish compartments are possible. When more compartments are present, more rate constants are needed to describe the kinetic behaviour of the compound.

6.3.6 Mesocosm studies

Aquatic model ecosystems that are most likely to be used for chemical registration purposes are the micro- and mesocosms. Guidance on performance of mesocosm studies is given in a series of publications, e.g. OECD (2006b); Campbell et al. (1999); Hill et al. (1994); Arnold et al. (1991). The summary and evaluation of the field test should support the objective of risk assessment. For extensive guidance on summarising and evaluating mesocosm studies, the reader is referred to De Jong et al. (2008).

6.4 Bees

6.4.1 General aspects

Data on the toxicity to honey bees should be submitted when exposure is likely (see HTB). Most acute bee toxicity studies are performed according to EPPO Guideline 170 or OECD 213 and 214. Summary Table 6-4 is based on these guidelines. Larval toxicity, long-term residual effects and disorientating effects can only be evaluated when tunnel, cage or field experiments are available (see Section 6.4.3). When considering potential effects to pollinators, the effects to honey bees are generally taken into account. However, bumblebees could be affected as well, as might be expected due to the taxonomical similarity. It may be assumed that bumblebees have a comparable or lower susceptibility particularly due to a more favourable proportion of surface and weight. This premise is, however, not well validated. Moreover, bumblebees behave differently which may be of influence. Specific Guidelines for bumblebees are under development.

6.4.2 Bee studies according to BBA 23-1 (1980) Guidelines

Only the oral toxicity test of BBA 23-1 is comparable to the EPPO Guidelines. However, the temperature can be too low, the solvent control may be lacking, and the maximum number of concentrations (three) is too low for a proper dose-response relation. For these reasons this test might have a limited value. BBA tests concerning contact or inhalation toxicity, and exposure to spray aerosols or dispersed powder are not comparable to the EPPO Guideline. These tests should not be summarised and they are not used for the Risk assessment.

BBA-data sheets do not meet the desired standards for reporting experimental set-up and results. In principle, they are assigned Ri 3, unless the applicant can provide additional information.

Table 6-4 Bees, LD_{50} studies

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance	improperly characterised or reported?	Y [→Ri 3]
2. test species	preferably young adult worker bees; age and race should be known	Y
3. oral exposure test		
3.1 test solution	10 or 20 µL test solution, containing 20 – 50 % sucrose solution	E
3.2 starvation	maximally 2 h before test	Y
4. contact test		
4.1 method of anaesthetic	e.g. CO_2: application should not lower temperature in cage	E
4.2 way of application	volume not reported?	Y
5. application		
5.1 number of concentrations	minimal five for correct statistical interpretation	E
5.2 vehicle	acetone is preferred ; vehicle used: solvent control should be tested	E
6. duration of test	not reported?	Y
6.1 treatment time (LD_{50} studies)		
6.2 observation time	minimum 48 h, maximum 96 h	Y
7. feed type		
8. number of animals	min. 3 groups of 10 bees/concentration	E
9. housing	cages should not cause control mortality; cages should be ventilated well	Y
10. test conditions (temperature, RH, light)	25 ± 2 °C; RH 60 – 70 %	Y
11. reference substance (positive control)	e.g. parathion, dimethoate	Y
RESULTS		
1. mortality in control groups	should not exceed X % [X = 15]	Y
2. overt signs of toxicity	changes in behaviour	
3. LD_{50} (95% CI)	recalculate if necessary both oral and contact LD_{50}'s are relevant	
PAY ATTENTION TO		
1. method of collection of bees	collection in early spring or late autumn should be avoided	E
2. correction for control mortality		
3. housing	bees should not be confined for more than one hour	E

6.4.3 Bee brood tests and insect growth regulators

The HTB requires that bee brood tests to be performed for Insect Grotwh Regulators (IGR). IGR may affect the bee brood by disturbing ecdysis and growth. This may cause long term effects. Noted insect growth regulators are acylurea compounds such as diflubenzuron and teflubenzuron. Other regulators are fenoxycarb — hormone mimicking — and cyromazine. There are no standard guidelines for bee brood tests, the method of Oomen et al. (1992) may be used. Bee brood tests should be seen as a rough screening laboratory test with an intensive exposure. In case effects are found in such tests, cage or tunnel tests are indicated (P. Oomen, pers. comm.) and/or a field test according to OECD Guidance (OECD, 2007) may be performed (see also 6.4.3).

6.4.4 Bees (effects other than acute toxicity)

There are no particular protocols of tests aiming at other endpoints than acute mortality, although in some particular cases tests on the influence on antenna response have been submitted. Other possibilities are bee poisoning by contaminated nectar or pollen, delayed mortality due to slow pesticide release of e.g. micro-encapsulated formulations, or behavioural effects.
For the assessment of deviating tests, it is advised to consult a bee expert of the Institute of Applied Plant Research (PPO) at Wageningen University and Research Centre.

6.4.5 Reviews on acute toxicity

Based on expert advise, bee studies reviewed in Beran (1970), Kupetz et al. (1979), and Stevenson (1978) get Ri 1. Bee studies reviewed in Atkins et al. (1981) get Ri 2. The toxicity values in Beran (1970) are reported as µg/g bee. This value should be divided by 10 to obtain the amount of µg/bee, as one bees is assumed to weigh 0.1 gram

6.4.6 Field tests

Guidance on field testing with bees is given in OECD (2007). It should be noted that in the context of this report, all outdoor bee toxicity tests (e.g. the tunnel and the cage test) are considered to be field tests as the test conditions mimic semi-realistic or realistic outdoor conditions. Honeybees and bumblebees are generally not exposed via pesticide residues on leaves or stems: they are only topically exposed to the inner petals of flowers. It is therefore important that application is performed when bees are active.
Tunnel and cage tests are performed with relatively small hives. Generally mortality is scored or the population density is estimated, and flight intensity is scored. However, under these conditions, mortality due to the bee captivation itself may be high. Therefore, it is important to take proper controls into account, including a positive control. Exposure in tunnel and cage tests is generally more intense than in full-scale field tests.
Field tests are difficult to perform: one should be sure that the treatment groups are indeed exposed to the crop sprayed chemical. As bees from a particular hive may have a radius of action of c. 3 km, this notion should be taken into account when performing a field test, as other hives — not designated as treatment groups — within this range may interfere with the test results. A more general problem with field tests is their replication: the use of a positive control may be controversial or even forbidden. Then additional, extensive observation of e.g. bee-keepers during the field test may give important clues. The use of a negative control with a chemical of which the bee-safeness has already been demonstrated is recommended. Expert judgement is recommended for the interpretation of the results.

6.5 Non-target arthropods other than bees

6.5.1 Laboratory tests

Most studies that are currently submitted are performed according to the guidelines as included in Candolfi et al. (2000), which form the basis of Summary Table 6-5. Older studies may have been performed according to the guidelines as compiled in Hassan (1992) or according to individual EPPO guidelines. The *Encarsia* studies of Oomen et al. (1994) get Ri 1.

First-tier laboratory studies are performed on inert substrates (most often glass or quartz sand). When these studies indicate a potential risk, extended laboratory studies on natural substrate (i.e. sprayed leaves, seedlings, natural soil) can be performed. The effect of residue decline can be studied by introducing an ageing period (indoors or outdoors), before the test organisms are exposed to the substrate.
In all laboratory studies, the effect on survival is determined, and depending on the result of the initial exposure, the test can be extended with a reproduction assessment. For most species, this latter phase is performed on untreated substrate. For beetles, food consumption is also a sub-lethal parameter. The effect on survival is most often calculated according to Abbott's formula (Abbott, 1925), with improvements of Schneider-Orelli (1947):

$$M = \left[\frac{M_t - M_c}{100 - M_c} \right] * 100$$ Equation 17

in which:
M = mortality in the treated group(s), corrected for the mortality in the control [%]
M_t = mortality in the treated group(s) [%]
M_c = mortality in the control group(s) [%]

In older studies, the combined effect on mortality and reproduction, expressed as the effect on beneficial capacity according to Overmeer et al. (1982), is used as an evaluation criterion:

$$E = 100\% - [(100\% - M) \times \frac{R_t}{R_c}]$$ Equation 18

in which:
E = effect on the beneficial capacity[%]
M = mortality in the treated group(s), corrected for the mortality in the control [%] (Equation 17)
R_t = the extent of reproduction in the treated group(s)
R_c = the extent of reproduction in the control group(s)

R_t and R_c can be determined for various parameters, e.g. as the number of successful parasitations, the number of eggs per female. For beetles, consumption is used as a parameter.
It should be noted that the recent guidelines do not include such a combined assessment on beneficial capacity, and the effect on lethal and sub-lethal parameters is evaluated and reported separately.

Nowadays, most studies on inert substrate are performed according to a dose-response design, in order to estimate the lethal rate for 50 % of the organisms, the LR_{50}. Information on how to deal with 'old' and 'new' tests in the risk assessment is given in the HTB.
Extended laboratory tests are also often performed according to a dose-response design, in order to evaluate the regulatory trigger of 50 % effect on survival or sub-lethal parameters. Because both lethal and sub-lethal parameters should meet this trigger, the evaluator should check whether or not there is 50 % effect on sub-lethal parameters at the level of the LR_{50}. This applies for instance to a situation when the estimated LR_{50} is between two application rates, the lower of which shows < 50 % effect on reproduction and the higher > 50 % effect. In this case, it should be stated in the Remarks that it is not clear whether < 50 % effect on reproduction would have occurred at the level of the LR_{50}. It can then only be concluded that < 50 % effect on survival and reproduction occurred at the lower application rate.

In conformity with preferred notation of the List of Endpoints prepared within the context of Directive 91/414/EEC, the following applies: When in a study the mortality in a treatment group is lower than in the control, the control corrected mortality for this treatment is indicated with a '-'sign, e.g. -5 %. When for a sublethal parameter a stimulation if found, the effect percentage is given with a '+'sign. To avoid any confusion, the following footnote should be added to the header table: '*a – sign for mortality and/or a + sign for sublethal effects indicate a favourable effect as compared to the control*'.

Table 6-5 Non-target arthropods, LR_{50} or use rate studies

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance	improperly characterised or reported?	Y [→Ri 3]
2. test species	not reported?	Y [→Ri 3]
2.1 rearing or collection method	not reported? except for wolf spiders (*Pardosa* spec.), all accepted species are from laboratory breeding cultures	Y
2.2 age or life stage	not specified?	Y
3. test substrate	not reported?	Y [→Ri 3]
4. application		
4.1 number of application rates	LR_{50}: minimal five for correct statistical interpretation	E
4.2 spray volume	should reflect agricultural practice should be checked by calibration	E
4.3 drying time	not reported?	E
5. replication	see guideline	E
6. duration of test	not reported? exposure time generally shorter than test duration	Y
7. test conditions (temperature, RH, light, feeding)	check with guideline; not reported?	Y
8. reference substance (positive control)	not included and no effects of test substance? not included and effects of test substance?	Y [→Ri 3] E
RESULTS		
1. validity criteria for control groups	see guideline	Y
2. missing animals	should be included in mortality calculation	
3. LR_{50} (95% CI)	recalculate if necessary	
4. sub-lethal effects	relevant for extended laboratory studies in case of LR_{50} studies: < 50 % effect demonstrated at level of LR_{50}?	E
PAY ATTENTION TO		
1. influence of co-formulants/adjuvants	animals may suffer from mechanical effects rather than from toxicity	

6.5.2 Field tests

Field or semi-field tests with non-target arthropods may involve several types of experiments, e.g. bioassays in which laboratory bred animals are exposed to the test substance in an enclosure under field conditions, studies in which a field population is studied before and after treatment or studies in which the arthropod population or community in treated plots is compared to that in an untreated control. Guidance on summarising and evaluating field studies on arthropod communities is currently under development by the Platform Higher Tier Studies.

6.6 Earthworms

6.6.1 Acute and chronic toxicity

Note that the scientific name of the most commonly used test species was changed: *Eisenia foetida andrei* is *Eisenia andrei*; *Eisenia foetida foetida is Eisenia fetida*.

The risk assessment for earthworms is based on an acute test in soil according to OECD 207 or ISO 11268-1, chronic tests should be performed according to ISO 11268-2 (Summary Table 6-6). An LC_{50} value from a test with filter paper, or from an Artisol test (a medium of silica gel; Reinecke, 1992) is not useful for risk assessment. The chronic risk for earthworms is based on the NOEC for reproduction from a study according to OECD 222 or ISO 11268-2. In both the acute and chronic tests, effects on

body weight change are usually reported. This information should be included in the Results section, but since it is not used as a regulatory endpoint, the results are not presented in the Header.

The experimental soil can be a natural soil, or an artificial (laboratory composed) soil. An example of artificial soil test substrate (OECD 207): 10 % sphagnum peat, 20 % kaolin clay; 70 % industrial sand; calcium carbonate is added to adjust the pH to 6.0 ± 0.5. The water content should be 40 – 60 % of maximum Water Holding Capacity (MWHC), depending on the peat used this is around 35 % of dry weight soil. The amount of organic matter may be adjusted (lowered) to meet specific regulatory purposes.
In reproduction studies, the test substance is often applied as a surface spray. In this case, the endpoint is expressed in (k)g as/ha. Since the NOEC is based on initial application rates, this value may be compared directly with the use rate, but only in case of a single application. When the NOEC has to be compared with a PEC_{soil} in mg/kg, the application rate in the test can be converted to mg/kg using the amount of test substance actually applied to the soil and the weight of the test soil. If this information is not reported, Van der Linden et al. (2008) recommend to perform the conversion assuming homogeneous distribution over 5 cm soil depth and a soil bulk density of 1000-1700 kg/m[4]. If the depth of the soil layer in the test is less than 5 cm, the actual depth should be taken.
It should be noted, however, that in the opinion related to the revision of Annex II and III of Directive 91/414/EC, EFSA-PPR does not support tests with sprayed soil surfaces. EFSA-PPR states that '*the risk assessment should be based on concentrations, and it has not been demonstrated that both application methods will yield the same results. From the scarce literature available it appears that the method of application has influence on the response (Van Gestel et al., 1995). The PPR Panel is therefore concerned that the assessment based on overspray is not informative and the decision making between products will not be uniform. It is the view of the PPR Panel that testing should be performed with test substances homogenously incorporated into the soils.'*(EFSA, 2007a).

6.6.2 Field tests

Extensive guidance for summarising and evaluating earthworm field studies is given in De Jong et al. (2006).

[4] A value of 1000 kg/m^3 is conservative if a total content has to be established whereas a value of 1700 kg/m^3 is conservative for pore water concentrations.

Table 6-6 Earthworms, toxicity studies

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance	improperly characterised or reported?	Y [→Ri 3]
2. test species 2.1 age and weight	usually performed with *Eisenia fetida* or *E. andrei* adult, min. 2 months old with clitellum; individual wet weight should be 300-600 mg	 Y
3. test soil 3.1 soil type 3.2 characteristics	 constituents of OECD soil not reported? natural soil not specified? pH, moisture checked? pH 6.0 ± 0.5 and moisture content should be 40 – 60 % of WHC (ca. 35% of dry weight)	 Y Y Y
4. application 4.1 way of application 4.2 number of concentrations 4.3 replicates 4.4 vehicle	 spray or mixing X treatment levels (geometric series) [X= 5] limit test allowed if LC_{50} >1000 mg/kg soil min. X replicates [X = 4] with 10 worms each vehicle control should be included	 E Y Y
5. analysis method	analysis not compulsory if performed: should be appropriate for the substance and/or metabolites, and the recovery of the substance should be >X% [X = 70] and <Y% [Y = 110]	 E
6. duration of exposure and observation period 6.1 acute test 6.2 chronic test	 mortality is assessed 7 and 14 days after application mortality and weight of F0 determined after 28 d; number of juveniles after 56 d.	 E E
7. test conditions 7.1 acute test 7.2 chronic test	 continuous light; 20 ± 2 °C; no feeding 16:8 h L:D; 20 ± 2 °C; feeding	 E E
RESULTS		
1. mortality in control groups	mortality in the controls should not exceed X% at the end of either test [X= 10]	Y
2. LC_{50} (95% CI) or NOEC	LC_{50}: 14-days, 7-days value not used NOEC: 56 days # juveniles	
3. sub-lethal signs of toxicity		
3.1. body weight change	reported, endpoint not in header	
3.2 behaviour	lack of burrowing may influence exposure	E
PAY ATTENTION TO		
1. moisture content	35 % of dry weight is not the same as 35 % of WHC	E

6.7 Micro-organisms in soil (toxicity)

A specific feature of tests with micro-organisms is that the uptake of substances is in general very quick. Therefore the effect can be apparent after half an hour up to two hours after application. Also the adaptation can be very quick.

Three types of tests can be distinguished:
1. test on soil processes (e.g. respiration, nitrification according to OECD 216/217);
2. test on the activity of enzymes (e.g. dehydrogenase, phosphatase, arylsulphatase);
3. single species test (e.g. Microtox).

Summary Table 6-7 refers to the tests on soil processes that in general are the only tests on micro-organisms that are included in the dossier. In rare cases, tests on enzyme activity related to soil respiration, such as dehydrogenase, may be submitted as additional information. Sometimes tests are

carried out in which the effects on ammonification (organic-N to NH_4^+) or on denitrification (NO_3^- to N_2) are measured. The Microtox test is carried out with e.g. *Photobacterium phosphoreum*, a salt water bacteria. The inhibition of light production is measured. This test is not considered relevant for the risk assessment of microbial processes in soil.
Nitrification and respiration tests are generally performed with natural soil. The agricultural history of the soil (pesticide use, fertilisation) and storage conditions after sampling are important to judge whether the test is performed with viable microbial communities.

Table 6-7 Micro-organisms and enzymes in soil and manure: N-transfer

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance	improperly characterised or reported?	Y [→Ri 3]
2. test soil		
2.1 soil type	USDA-class and other relevant data reported? soil must be relevant for the agricultural situation (e.g. no paddy or volcanic soil; 0.5 – 15 % OM)	Y
2.2 characteristics	pH, CEC, OM reported?	Y
2.3 moisture content	not reported?	Y
3. application		E
3.1 number of concentrations	min. X concentrations [X = 2]: the recommended dose, and 10 times the recommended dose	E
3.2 replicates	min. X replicates [X = 3]	Y
3.3 vehicle	vehicle control should be included	Y
4. additives (lucerne meal, ammonia)	Micro-organisms can be influenced negatively by too high concentrations of some additives (e.g. ammonium sulphate)	E
5. sampling frequency	min. X samples [X = 3]: after 7, 14 and 28 days	E
6. analysis method for N	not reported?	E
7. conditions	temperature should be X °C [X = 15 - 25] dark conditions are preferred	E
RESULTS		
1. conversion rate for each time interval	rates should be compared, not concentrations; if necessary, rates should be calculated	
2. maximum % effect on conversion rate relative to control over time window	report time point	
3. % effect at end of test	< 25 %, otherwise test should be prolonged	Y
PAY ATTENTION TO		
1. agricultural history of soil	no manipulation with fertiliser, no (prior) use of pesticides that may have lead to adapted micro-organisms (in the previous five years). Special attention should be paid to compounds interfering with the N-cycle in the soil.	E
2. storage	if there is no immediate use, storage in the lab or in the open should be appropriate (see 5.1.1)	Y

6.7.1 Interpretation of nitrogen transfer studies (OECD 216) for risk assessment

This section deals with the interpretation of nitrogen transfer studies for risk assessment of plant protection products within the context of Directive 91/414/EEC, and strictly speaking it is thus partly beyond the focus of this manual (see also 2.1 and 2.2). The endpoint to be derived from these studies is, however, based on the regulatory trigger, and this is reflected in the way these studies are summarised and evaluated.
Nitrification is a process in which several species of micro-organisms are involved. The process consists of the oxidation of ammonium to nitrite and the subsequent oxidation of nitrite to nitrate. The process of nitrification is relatively susceptible to disturbance. In the nitrogen transformation test, the conversion of nitrogen (added as lucerne meal) to nitrate is studied. The interpretation of these studies is sometimes difficult. The following problems may arise:

- in the report, the *concentrations* of nitrate (nitrite and ammonium) after 28 days are compared, while according to OECD 216 *formation rates* (i.e. changes in concentration over time) should be evaluated;
- non-significant differences in concentrations between control and treatment can lead to a non-significant difference in formation rate of > 25 % (i.e. regulatory trigger)
- some authors base the evaluation on nitrate, others on mineral N-content (sum of nitrate, nitrite and ammonium);
- temporary decreases in nitrate content may occur;
- control and treatment differ at the start of the test with respect to N-content.

These problems are mainly due to the fact that the quality of the lucerne meal and the characteristics and treatment of the soil may highly influence the performance of the test. In addition, OECD 216 allows for a 15 % variation between the control replicates (it is assumed that coefficient of variation, CV, is meant, i.e. the standard deviation as percentage of the mean). With the latter criterion, non-significant differences in formation rate of > 25 % can easily occur.

The following aspects are used as a guidance for interpretation and evaluation of the test results (based on J.W. Vonk, EPP Consultancy, pers. comm.):

- If the CV in the *control* is > 15 %, the test gets Ri 3.
- If at the start of the test the variation between the replicate treatment samples is > 15 %, this may indicate that the test is not properly carried out. If, however, in this case the formation rate after 28 days is ≤ 25 % significantly different from the control, it is considered to be demonstrated that the regulatory trigger of ≤ 25 % effect after 100 days is met.
- The nitrate concentration in the control soil should show a steady increase with time, a decrease is often indicative of denitrification as a result of anaerobic conditions due to e.g. insufficient aeration. Net nitrate fixation can, however, occur for unknown reasons. If a decrease is observed only once during the 28-days test duration, this is considered acceptable.
- If, considering the above, a test is considered reliable (i.e. CV in control and treatments < 15 %), the nitrate concentrations can be plotted against time to evaluate whether or not > 25 % effect on nitrate formation rate is present over the test duration.
    - If no clear trend towards inhibition or stimulation is present, and the concentrations of nitrate after 28 days differ by ≤ 25 % from the control, the regulatory trigger is considered to be met, because in that case the formation rate is also different by ≤ 25 % as compared to the control.
    - If the concentrations of nitrate after 28 days differ by > 25 %, but the differences are not significant, then the trigger is also considered to be met.
    - If the concentrations of nitrate after 28 days differ by > 25 %, and the difference are significant, the test should be prolonged or a new test, longer than 28 days, should be performed.
    - If there is a trend towards inhibition or stimulation during the first 14 days, and the formation rate over 14 – 28 days differ by > 25 % (significant or non-significant), the test should be prolonged or a new test longer than 28 days should be performed.
- According to OECD 216, only nitrate has to be analysed, but according to the Uniform Principles, nitrate mineralisation should be evaluated (i.e. ammonium + nitrate + nitrite). Both are accepted.
- If accumulation of ammonium occurs, this may be indicative of an inhibition of the nitrification step.
- Nitrite may be detected when the 2nd stage of nitrification process has not been built up yet. This phenomenon should only be present temporarily. If a lot of nitrite is formed in the control, this is indicative of a bad performance of the test. If nitrite is formed only in the test item treatment, this is indicative of an effect of the test substance.

- Deviating test results often occur in soils with a low nitrogen mineralisation capacity. Since nitrogen mineralisation capacity on one location can be highly variable, it is not possible to set a criterion for this aspect.

6.8 Respiration in activated sludge

Tests on the effects of substances on the respiration of bacteria in activated sludge are performed to assess the potential impact upon emission to a sewage treatment plant. This is relevant for a number of biocides and for certain uses of agricultural pesticides, a.o. mushroom cultivation, flower bulb or potato treatment and substrate culture. Tests can be performed according to OECD 209.

Table 6-8 Influence on activated sludge (respiration)

ITEMS	NOTES	RELIABILITY LOWER?
DESCRIPTION		
1. test substance	improperly characterised or reported?	Y [→Ri 3]
2. test system	improperly characterised or reported	E
2.1 type and source of microbial inoculum	usually activated sludge from a municipal or an industrial sewage treatment plant (STP)	
2.2 adaptation	sludge should not be adapted to the test substance	Y
2.3 nutrient solution		
2.4 pH	not reported?	Y
3. application		E
3.1 number of concentrations	at least X concentrations [X= 5] should be tested; difference between concentrations should not exceed factor 3.2	E
3.2 replicates	min. X replicates [X = 2]	Y
3.3 vehicle	vehicle control should be included	Y
4. conditions: temperature, aeration	aeration should take place	Y
5. reference substances	at least three concentrations of 3,5-dichlorophenol	Y
RESULTS		
1. control variation	the two control respiration rates should not deviate by 15% of each other	Y
2. EC_{50} (95% CI)		
2.1 EC_{50} reference substance	3-hours EC_{50} of 3,5-dichlorophenol should be in range 5 - 30 mg/L	Y
PAY ATTENTION TO		
1. water solubility	test concentrations should not exceed water solubility	E
2. the dissipation type	there should be no major loss due to hydrolysis, photolysis, or volatilisation: is the test design adequate?	E

References

Abbott WS. 1925. A method of computing the effectiveness of an insecticide. Journal Econ. Entomol. 18: 265-267.

Anderson T-H and Domsch KH. 1989. Ratios of microbial biomass carbon to total organic carbon in arable soils. Soil Biol Biochem 21: 471-479.

Arnold D, Hill I, Matthiesen P, Stephenson R. 1991. Guidance document on testing procedures for pesticides in freshwater static mesocosms - from the workshop 'A meeting of experts on guidelines for static field mesocosm tests', Monks Wood Experimental Station, Abbotts Ripton, Huntingdon, UK. Brussels, Belgium: SETAC-Europe. 46 pp.

Atkins EL, Kellum D, Atkins KW. 1981. The reducing of pesticide hazards to honey bees: mortality prediction technics and integrated management strategies. University of California, Division of Agricultural Sciences. Leaflet 2883, p.23.

BBA. 1986. Richtlinien für die amtliche Prüfung von Pflanzenschutzmitteln. Teil IV, 4-2: Versickerungsverhalten von Pflanzenschutzmitteln (vormals BBA-Merkblatt Nr. 37).

Beran F. 1970. Der gegenwärtigen Stand unserer Kenntnisse über die Bienengiftigkeit und Bienengefährlichkeit unserer Pflanzenschutzmittel. Gesunde Pflanzen 22: 21-31.

Campbell PJ, Arnold DJS, Brock TCM, Grandy NJ, Heger W, Heimbach F, Maund SJ, Streloke M, Eds. 1999. Guidance document on higher-tier aquatic risk assessment for pesticides (HARAP), from the SETAC-Europe/OECD/EC Workshop. 19-22 April 1998 Lacanau Océan, France. Brussels, Belgium: SETAC-Europe. 179 pp.

Candolfi, MP, Blümel, S, Forster, R, Bakker, FM, Grimm, C, Hassan, SA, Heimbach, U, Mead-Briggs, MA, Reber, B, Schmuck, R, and Vogt, H. 2000. IOBC, Bart and EPPO Joint Initiative. Guidelines to evaluate side-effects of plant protection products to non-target arthropods.

De Jong FMW, Brock TCM, Foekema EM, Leeuwangh P. 2008. Guidance for summarizing and evaluating aquatic micro- and mesocosm studies. Bilthoven, The Netherlands: Report no. 601506009. 51 pp.

De Jong FMW, Van Beelen P, Smit CE, Montforts MHMM. 2006. Guidance for summarising earthworm field studies. Bilthoven, The Netherlands: RIVM. Report no. 601506006. 46 pp.

EC. 2002a. Guidance document on risk assessment for birds and mammals under Council Directive 91/414/EEC. Brussels, Belgium: Health & Consumer Protection Directorate General. Report no. SANCO/4145/2000-final. 43 (with Annexes) pp.

EC. 2002b. Working Document. Guidance document on aquatic ecotoxicology in the context of the Directive 91/414/EEC. Brussels, Belgium: Health & Consumer Protection Directorate General. Report no. SANCO/3268/2001 rev.4 (final). 61 (with Annexes) pp.

EC. 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 and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Ispra, Italy: European Commission Joint Research Centre.

EFSA. 2007a. Opinion of the Scientific Panel on Plant protection products and their Residues on a request from the Commission related to the revision of Annexes II and III to Council Directive 91/414/EEC concerning the placing of plant protection products on the market - Ecotoxicological studies. The EFSA Journal 461: 1-44.

EFSA. 2007b. Opinion on a request from EFSA related to the default $Q10$ value used to describe the temperature effect on transformation rates of pesticides in soil. The EFSA Journal 622: 1-32.

FOCUS. 2006. Guidance document on estimating persistence and degradation kinetics from environmental fate studies in EU registration. The final report of the work group on degradation kinetics of FOCUS (Forum for the Co-ordination of pesticide fate models and their USe). Brussels, Belgium: EC/DG SANCO. Report no. EC Document Reference SANCO/10058/2005, version 2.0. 434 pp.

Gobas FAPC and Zhang X. 1992. Measuring bioconcentration factors and rate constants of chemicals in aquatic organisms under conditions of variable water concentrations and short exposure time. Chemosphere 25: 1961-1971.

Haanstra L, Doelman P, Oude Voshaar JM. 1985. The use of sigmoidal dose response curves in soil ecotoxicology. Plant Soil 84: 293-297.

Hamilton MA, Russo RC, Thurston RV. 1977. Trimmed Spearman-Kärber method for estimating median lethal concentrations in toxicity bioassays. Environ Sci Technol 11: 714-719. Correction 1978, vol. 12: 417.

Hassan SA. 1992. Guidelines for testing the effects of pesticides on beneficial organisms: Description of test methods. IOBC/WPRS Bulletin XV/3

Hill IR, Heimbach F, Leeuwangh P, Matthiessen P, Eds. 1994. Freshwater field tests for hazard assessment of chemicals. Ann Arbor, USA: Lewis Publishers. 561 pp.

Kupetz H, Nieder G, Wegner H. 1979. Ergebnisse von Untersuchungen zur Bienengefährlichkeit und Bienentoxizität von Pflanzenschutzmittel und deren Wirkstoffen. Pflanzenschutzberichte 45: 165-188.

Luttik R. 1993. Environmental hazard/risk assessment of pesticides used in agriculture for birds and mammals: the Dutch concept. Part 2. Avian food avoidance behaviour. Bilthoven, The Netherlands: RIVM. Report no. 679102019.

Mensink BJWG, Montforts M, Wijkhuizen-Maslankiewicz L, Tibosch H, Linders JBHJ. 1995. Manual for summarising and evaluating the environmental aspects of pesticides; with Addendum August 1996. Bilthoven, The Netherlands: RIVM. Report no. 679101022. 117 pp.

OECD. 2006a. OECD Series on Testing and Assessment. 54: Current approaches in the statistical analysis of ecotoxicity data. A guidance to application.

OECD. 2006b. OECD Series on Testing and Assessment. 53: Guidance document on simulated freshwater lentic field tests (outdoor microcosms and mesocosms).

OECD. 2007. OECD Series on Testing and Assessment. 75: Guidance document on the honey bee (Apis mellifera L.) brood test under semi-field conditions.

Oomen PA, De Ruijter A, Van der Steen J. 1992. method for honeybee brood feeding tests with insect growth-regulating insecticides. Bulletin OEPP/EPPO Bulletin 22: 203-216.

Oomen PA, Jobsen JA, Romeijn G, Wiegers GL. 1994. Side-effects of 107 pesticides on the whitefly parasitoid *Encarsia formosa*, studied and evaluated according to EPPO Guideline no. 142. EPPO Bulletin 24: 89-107.

Overmeer WPJ, Van Zon AQ, Adriaanse PI. 1982. A standard method for testing the side effects of pesticides on the predacious mite, *Amblyseius potentilla* (Acarina: Phytoseiidae). Entomophaga 27: 357-364.

Reinecke AJ. 1992. A review of ecotoxicological test methods using earthworms. In: Greig-Smith, PW, Becker, H, Edwards, PJ, Heimbach, F, Eds. Ecotoxicology of earthworms,

Schneider-Orelli O. 1947. Entomologisches Praktikum . Aarau , 2. Auflage.

Stevenson JH. 1978. The acute toxicity of unformulated pesticides to worker honey bees (*Apis mellifera* L.). Pl Path 27: 38-40.

USDA. 1951. Soil Survey Manual. U.S. Dept. Agriculture Handbook No. 18. Washington, USA: Agricultural Research Administration, United States Department of Agriculture.

Vaal MA, Van den Dikkenberg RPh, Van de Plassche EJ. 1992. Poorly water soluble organic compounds: Guidance for the interpretation of results from aquatic ecotoxicity tests and a proposal for a modified classification and risk evaluation. Bilthoven, The Netherlands: RIVM. Report no. 679102016. 23 pp.

Van der Linden AMA, Boesten JJTI, Brock TCM, Van Eekelen GMA, Ter Horst MMS, De Jong FMW, Montforts MHMM, Pol. J.W. 2008. Revised proposal for the risk assessment of persistence of plant protection products in soil. Bilthoven, the Netherlands: RIVM. Report no. 601712003. 105 pp.

Van Ewijk PA, Hoekstra JA. 1993. Calculation of the EC_{50} and its confidence interval when subtoxic stimulus is present. Ecotoxicol Environ Saf 25: 25-32.

Van Gestel CAM. 1986. Inleiding tot de bodemkunde; toepassing bij het beoordelen van gegevens over het gedrag van stoffen in de bodem. Bilthoven, The Netherlands: RIVM. Report no. 718600001.

Van Gestel CAM, Zaal J, Dirven-Van Breemen EM, Baerselman R. 1995. Comparison of two test methods for determining the effects of pesticides on earthworm reproduction. Acta Zool Fenn 196: 278-283.

Van Vlaardingen PLA and Smit CE. 2008. Verification of moisture content in test soils. Bilthoven, The Netherlands: RIVM. Report no. 601516016.

Verschoor AJ, Boesten JJTI, Leistra M, Van der Linden AMA, Linders JBHJ, Pol JWW. 2001. Evaluation of pesticide leaching in lysimeter and field studies. Parent substances. Bilthoven, The Netherlands: RIVM. Report no. 601506007. 44 pp.

Whitehouse P and Mallet M. 1993. Aquatic toxicity testing for notification of new substances. An advisory document on dealing with ‘difficult substances’. Report to the Chemical Notification Unit, Department of the Environment. Medmenham, United Kingdom : WRC. Report no. CP 722. 34 pp.

rivm

Annex 1 Useful formulas, units, conversions and air saturation values

Latin prefixes

deci	(d)	$= 10^{-1}$	deca	(da)	$= 10^{1}$
centi	(c)	$= 10^{-2}$	hecto	(h)	$= 10^{2}$
milli	(m)	$= 10^{-3}$	kilo	(k)	$= 10^{3}$
micro	(μ or u)	$= 10^{-6}$	mega	(M)	$= 10^{6}$
nano	(n)	$= 10^{-9}$	giga	(G)	$= 10^{9}$
pico	(p)	$= 10^{-12}$	tera	(T)	$= 10^{12}$

Pressure

1 mm Hg	= 133 Pa [N/m^2]
1 mm Hg	= 1 Torr
760 mm Hg	= 101300 Pa
101300 Pa	= 1 atm
100 kPa	= 1 bar
1 psi (pounds per square inch)	= 6860 Pa

Soil

CEC: 1 meq/100 g = 10 mmol/kg

Water hardness

German degree (dH):	1° hardness =	17 mg/L as $CaCO_3$,
		10 mg/L as CaO
French degree:	1° hardness =	10 mg/L as $CaCO_3$
English degree:	1° hardness =	1 grain/gallon as $CaCO_3$
		10 mg/L as $CaCO_3$
American degree:	1° hardness =	1 mg/L as $CaCO_3$

Toxicity

TLm (median tolerance limit) is the LC_{50}. MATC is the mean of the NOEC and LOEC.

Light

1 footcandle = 10.76 lux

Conversion of English/American units

Note that various unit converters can be found on the internet

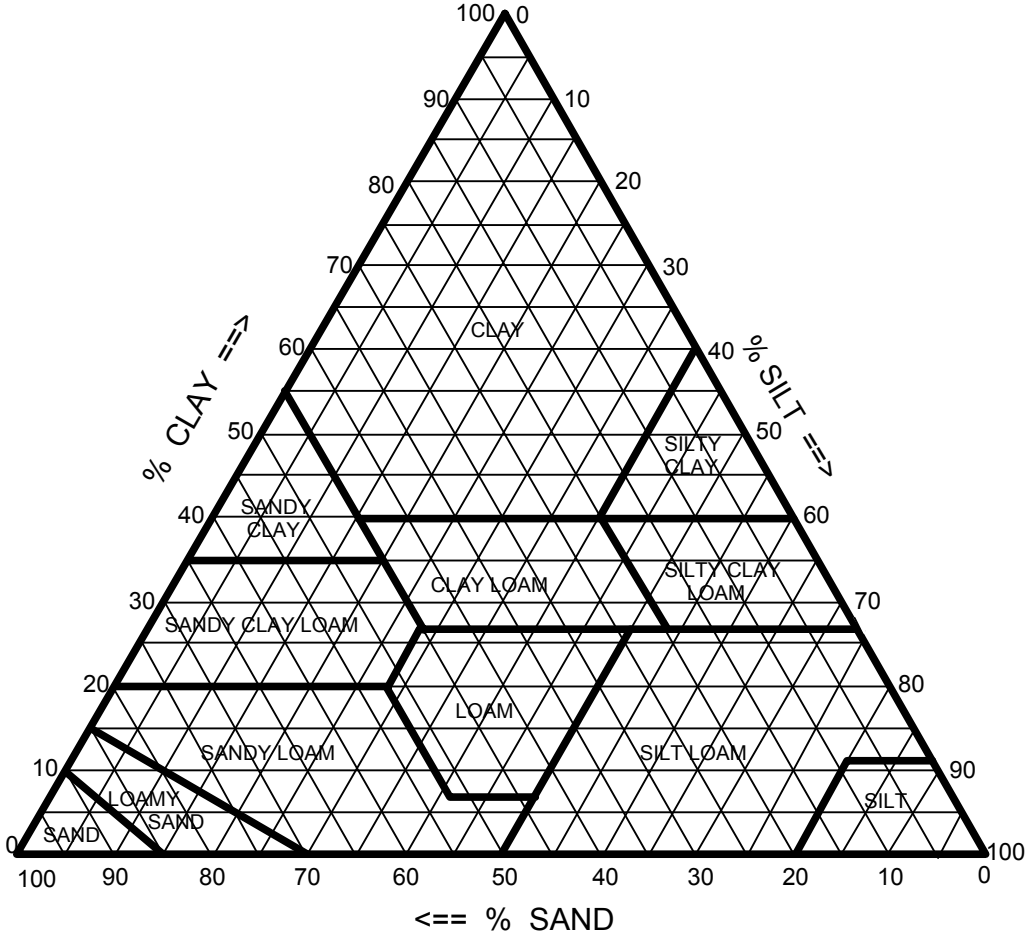
Light	Temperature
1 footcandle = 10.76 lux	1 °F = -17.2 °C
Length	**Volume**
1 inch = 2.54 cm	1 cubic inch = 16.3871 cm^3
1 foot= 12 inches = 30.48 cm	1 cubic foot = 28.3168 dm^3
1 yard = 3 feet = 0.9144 m	1 cubic yard = 0.76455 m^3
1 mile = 1.60934 km	1 pint = 1/8 engl. gallon = 0.568261 dm^3
	1 quart = 1/4 engl. gallon = 1.13652 dm^3
	1 engl. gallon = 4.54609 dm^3
	1 amer. gallon = 3.785 dm^3
	1 fluid pint = 1/8 amer. gallon
	1 fluid quart = 1/4 amer. gallon
Area	**Weight**
1 sq. inch = 6.4516 cm^2	1 grain = 64.7989 mg
1 sq. foot = 9.290304 dm^2	1 ounce (oz) = 28.35 g
1 sq. yard = 0.8361 m^2	1 pound (lb) = 0.453592 kg
1 sq. mile = 2.59 km^2	
1 acre = 4047 m^2	

O_2 saturation in water

temperature [°C]	Solubility of oxygen in mg/L (100 % saturation at 1 atm)				
	fresh water	seawater 5 g Cl/L	seawater 10 g Cl/L	seawater 15 g Cl/L	seawater 20 g Cl/L
10	11.3	10.7	10.1	9.6	9.0
11	11.1	10.5	9.9	9.4	8.8
12	10.8	10.3	9.7	9.2	8.6
13	10.6	10.1	9.5	9.0	8.5
14	10.4	9.9	9.3	8.8	8.3
15	10.2	9.7	9.1	8.6	8.1
16	10.0	9.5	9.0	8.5	8.0
17	9.7	9.3	8.8	8.3	7.8
18	9.5	9.1	8.6	8.2	7.7
19	9.4	8.9	8.5	8.0	7.6
20	9.2	8.7	8.3	7.9	7.4
21	9.0	8.6	8.1	7.7	7.3
22	8.8	8.4	8.0	7.6	7.1
23	8.7	8.3	7.9	7.4	7.0
24	8.5	8.1	7.7	7.3	6.9
25	8.4	8.0	7.6	7.2	6.7
26	8.2	7.8	7.4	7.0	6.6
27	8.1	7.7	7.3	6.9	6.5
28	7.9	7.5	7.1	6.8	6.4
29	7.8	7.4	7.0	6.6	6.3
30	7.6	7.3	6.9	6.5	6.1

Annex 2 Soil classification

Textural classes of mineral soils according to the US soil classification. Particle sizes: clay <2 µm, silt 2 - 50 µm, sand >50 µm. USDA (1951).



Annex 3 Test guidelines

Various (inter)national organisations are active in the field of test guideline development. Information on test guidelines can be obtained via the following links. Please note that URLs are subject to frequent change. In case the URLs listed below do not function correctly, information may be obtained via the search engines on the organisation's home pages.

OECD

The Guidelines for the testing of chemicals published by the Organization of Economic Co-operation and Development (OECD) are a collection of the most relevant internationally agreed test methods used by government, industry and independent laboratories to determine the safety of chemicals and chemical preparations, including pesticides and industrial chemicals. They cover tests for the physical-chemical properties of chemicals, human health effects, environmental effects, and degradation and accumulation in the environment. These guidelines are most often used in dossiers, because of their wide acceptance.
General information:
http://www.oecd.org/department/0,3355,en_2649_34377_1_1_1_1_1,00.html
Guidelines:
http://puck.sourceoecd.org/vl=2295854/cl=19/nw=1/rpsv/periodical/p15_about.htm?jnlissn=1607310x
In some cases, draft guidelines submitted to the OECD test guidelines program are not accepted because of limited validation or for other reasons. In that case, the test protocols are sometimes published as Guidance documents in the Series on Testing and Assessment:
http://www.oecd.org/document/30/0,3343,en_2649_34377_1916638_1_1_1_1,00.html

US-EPA

The United States Environmental Protection Agency developed a large number of guidelines, a key to which can be found via
http://www.epa.gov/epahome/index/

ISO

The International Organization for Standardization produced international standards on a variety of subjects, among which the assessment of environmental aspects of chemicals. Standards published by TC 147 (water quality) and 190 (Soil quality) are most relevant, some of these were transferred into OECD Guidelines. ISO-guidelines are available as pdf, but should be purchased. The catalogue is available via
http://www.iso.org/iso/iso_catalogue.htm

NEN

The Netherlands Normalisatie-instituut (NEN) is the Dutch member of ISO. The NEN has issued a number of ISO-standards as NEN-standard, but there are also some standards that are developed solely on a national scale. More information via
http://www2.nen.nl/nen/servlet/dispatcher.Dispatcher?id=231695

Environment Canada

The Method Development and Applications Section (MDAS) of Environment Canada develops test guidelines to meet the requirements of the Environment Canada programs. Some methods are adapted from OECD methods, but there are also methods for organisms and specific situations that are not covered by OECD.
http://www.etc-cte.ec.gc.ca/organization/bmd/bmd_publist_e.html

BBA

The German Federal Biological Research Centre for Agriculture and Forestry (Biologische Bundesanstalt, BBA) developed guidelines for the testing of agricultural chemicals. These so-called 'Richtlinien für die amtliche Prüfung von Pflanzenschutzmitteln' are published and sold via:
http://www.saphirverlag.de/
Some guidelines were previously published as 'Merkblätter'. More information, also on other subjects, can be found at the BBA's website, via
http://www.bba.bund.de/cln_044/nn_805654/DE/veroeff/veroeff__node.html__nnn=true

EU Guidelines

In the context of EU Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging, and labelling of dangerous substances, Annex V to Directive 84/449/EEC contains a description of test methods.
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31984L0449:EN:HTML

RIVM

National Institute
for Public Health
and the Environment

P.O. Box 1
3720 BA Bilthoven
The Netherlands
www.rivm.com