National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Guidance for the derivation of environmental risk limits

Part 3. Derivation of ERLs for freshwater and marine water

version 1.0

Colophon

Contact:
Helpdesk Risico's van Stoffen

This investigation has been performed by order and for the account of Ministry of Infrastructure and the Environment, within the framework of the project 'Nationaal stoffenbeleid ZZS'.

This is a publication of:
National Institute for Public Health and the Environment
P.O. Box 1 | 3720 BA Bilthoven
The Netherlands
www.rivm.nl/en

Contents

1 Introduction

1.1 Relationship with other guidance

As indicated in the chapter 'Introduction and definitions' (see ERL Report 01), the European guidance for derivation of environmental quality standards (EQS) within the Water Framework Directive [1] is followed for derivation of water quality standards. The present ERL-report only deals with those aspects that are not (fully) covered in the WFD-guidance and/or for which specific (national) guidance has been developed. Examples are the derivation of the additional national risk limits Serious Risk Concentration (SRC) and the Negligible Concentration (NC) (see ERL Report 01, section 3.3), and additional guidance developed in the context of a Dutch project on aquatic effects assessment [2,3]. Scientific developments in other regulatory frameworks, such as REACH, and authorisation of biocides and plant protection products are taken into account as well.

1.2 Terminology: EQS instead of MPC

When adopting the WFD-guidance it was decided to use the terminology of the WFD. As a consequence, the 'MPC' will no longer be used for water. For the water compartment, the WFD distinguishes two quality standards to cover both long- and short-term effects resulting from exposure: a long-term standard, indicated as the annual average environmental quality standard (AA-EQS) and normally based on chronic toxicity data, and a short-term standard, referred to as a maximum acceptable concentration EQS (MAC-EQS) which is based on acute toxicity data. In addition, a quality standard is derived for surface water that is used for drinking water abstraction. This is the concentration in surface water that meets the requirements for use of surface water for drinking water production. The terms AA-EQS and MAC-EQS are used in the European priority substances directive 2013/39/EU[1]. The Dutch equivalents of these terms are used in Dutch legislation based on the WFD and EQS is used throughout this report as well. The abbreviations refer to quality standards. However, the values that are derived based on the this guidance have a status of scientific advisory values and will be effective as official standards only after approval by the responsible ministries. This status of the results should be made clear when publishing reports based on this guidance.

1.3 Triggers for secondary poisoning and human health

The long-term WFD-water quality standards take account of direct ecotoxicity, but also include exposure of humans and predatory birds or mammals via intake of fish or shellfish. Inclusion of the latter routes depends on the characteristics of the compound with respect to bioaccumulation and human toxicology.

[1] Richtlijn 2013/39/EU van het Europees Parlement en de Raad van 12 augustus 2013 tot wijziging van Richtlijn 2000/60/EG en Richtlijn 2008/105/EG wat betreft prioritaire stoffen op het gebied van het waterbeleid.

1.3.1 Secondary poisoning

1.3.1.1 Organic compounds

Secondary poisoning should be assessed if there is evidence of bioaccumulation potential or the substance has a high intrinsic toxicity to mammals and birds (except carcinogenicity) [1]. Evidence of bioaccumulation potential is demonstrated according to the following criteria for the bioconcentration factor (BCF), bioaccumulation factor (BAF) and/or biomagnification factor (BMF) or log K_{ow}:

- the measured BCF or BAF ≥ 100 L/kg or BMF > 1, or
- no valid measured BCF (BAF) is available, but log K_{ow} ≥ 3, or
- if there is other evidence of bioaccumulation potential (e.g. biota monitoring data, structural alerts)
- PROVIDED THAT there is no mitigating property such as rapid degradation (ready biodegradability or hydrolysis half-life <12 h at pH 5-9, 20°C) or obvious molecular size exclusion

Detailed information on the collection, evaluation and selection of bioaccumulation data is given in section 3.1. When there is reason for concern because of a relatively high toxicity for birds or mammals, and/or bioaccumulative properties, it is advised to perform initial calculations to check whether secondary poisoning might be critical for EQS-derivation. If adequate bioaccumulation data are absent, it is worthwhile to back-calculate the BAF-value which would lead to a QS that is lower than that for direct ecotoxicity, and to evaluate whether or not this BAF is realistic in view of the characteristics of the compound (e.g. log K_{ow}, QSAR estimates of BCF). If relevant bird of mammal toxicity data are absent, but the BCF/BAF/BMF meet the trigger(s), an initial assessment on the basis of the human toxicological threshold limit (TTL_{hh}) may reveal if secondary poisoning might potentially be critical. The TTL_{hh} usually is derived on the basis of mammal data, but taking different endpoints into account. Such an initial assessment can be helpful to decide on the need for additional actions to retrieve bird and mammal data.

1.3.1.2 Metals

Text modified from WFD guidance, section 2.4.3.1, which should be consulted as well.

Biomagnification of metals in aquatic organisms is rarely observed and, if it does occur, it usually involves the organo-metallic forms of metals, e.g. methylmercury [4]. This should not be interpreted as lack of exposure or no concern for trophic transfer. Even in the absence of biomagnification, aquatic organisms can bioaccumulate relatively large amounts of metals and this can become a significant source of dietary metal to predators, including humans.

Therefore, the assessor should examine the potential of a metal to biomagnify or cause secondary poisoning in food. Review the information available for the metal in question in order to assess whether an in-depth secondary poisoning assessment is needed. Pay attention to possible formation of organometallic species in compartments.

BCF of BAF values cannot be used as such and cannot be compared to trigger values, as with organic compounds. For a number of metals an inverse relationship between BCF and external (water-) concentration is observed [5]. Consequently, BCFs and BAFs are not constant with water concentration. Furthermore, some metals are essential for life and many organisms possess mechanisms for regulating internal concentrations.

A case-by-case evaluation of the possibility of dietary toxicity is required and the following information needs to be considered:

- Information on metal mode of action and homeostatic (internal regulation) controls,
- Information on essentiality,
- Information on biomagnification (BMF). An relevant example study is that by Ikemoto et al. [6].
- Information on major toxicities i.e. whether main risks are through direct toxicity to pelagic organisms or secondary poisoning. With regards to the potential for secondary poisoning the assessment of the mode of toxic action in both prey and predator is a key consideration. If there is no evidence of biomagnification (i.e. BMF<1) and no specific toxicity in birds and mammals compared to fish (on a dose based approach), the $QS_{fw, eco}$ and $QS_{sw, eco}$ should be protective for birds and mammals as well as pelagic organisms.
- However, it must be realized that such information for dietary exposure of fish is often not available. In such case, a similar approach as described above for organic substances to assess what would be the critical BAF at $QS_{fw, eco}$ and $QS_{sw, eco}$ would give an initial impression whether secondary poisoning might be critical.

1.3.2 Human health

To decide on the inclusion of human health in risk limit derivation, information on the classification of a compound with respect to human toxicology is necessary. In short, a human health based standard should be derived for all substances which are classified as being (suspected) carcinogenic, reprotoxic of mutagenic (CMR) under the European regulation on Classification, Labelling and Packaging 1272/2008, irrespective of their bioaccumulation potential. This involves substances with H-statements H340, 341, 350, 351, 360, 361, 362. Human exposure should also be taken into account in case of classification with respect to danger of serious damage to health after prolonged exposure (H373), or oral toxicity (H300, 301, 302), in combination with the properties as described above for secondary poisoning.

Classification regarding (suspected) carcinogenicity may be obtained from tests with a less relevant exposure route, e.g. inhalation or dermal contact. It is assumed that carcinogenicity is an intrinsic property of the compound and unless there is clear evidence that dietary exposure does not result in cancer, inclusion of human exposure via fish is triggered. In addition, information obtained from evaluation of toxicological data that (as yet) is not necessarily reflected in classification and labelling phrases should be considered.

Similar to what is described above for secondary poisoning, it is advised to perform initial calculations to check whether human consumption of fish might be critical for EQS-derivation in case of a relatively low TTL_{hh} and/or bioaccumulative properties.

1.4 Terminology: QS and EQS

For each of the routes considered for EQS-derivation, the intermediate standards that are derived during the process of standard derivation are indicated as 'QS'. The lowest value is taken forward as final AA-EQS. The short-term standard MAC-EQS only considers direct ecotoxicity and strictly speaking an intermediate standard for different exposure routes is not needed. However, for reasons of consistency, the same terminology is used as for the long-term standard, using the term 'MAC-QS' during the process of derivation and 'MAC-EQS' for the final proposed standard. Subscripts are used to distinguish between the routes considered (eco: ecotoxicity, secpois: secondary poisoning, hh: human health) and the medium which the standard refers to (biota, fw: freshwater, sw: saltwater). The subscript 'water' is used for standards that are applicable to both fresh- and saltwater. Table 1 summarises the water quality standards and their respective abbreviations.

Separate EQSs are derived for freshwaters and saltwaters, which according to the WFD-guidance is justified by the need to account for the additional uncertainties associated with extrapolation for the marine ecosystem, especially the general under-representation in the experimental dataset of specific marine key taxa and possibly a greater species diversity. Transitional (e.g. estuarine) waters are intermediate in salinity which can vary on a diurnal cycle. According to the WFD-guidance, the freshwater scheme may be more appropriate for estuarine waters with a low salinity that are supporting communities that are closely related to freshwater ecosystems. Salinity levels between 3 and 5 ‰ can be considered as a switch from communities that are dominated by freshwater species to communities that are dominated by saltwater species. Therefore, the WFD-EQS are not reported for 'transitional ánd marine waters', but either for freshwaters or saltwaters. The WFD-guidance recommends a salinity of 5 ‰ as the cut-off unless other evidence suggests a different cut-off is appropriate for a particular location.

Table 1 Overview of the different types of quality standards for fresh- and saltwater according to the WFD-guidance. An asterisk indicates that considering this aspect depends on the characteristics of the compound.

Type of QS	**Protection aim**	**Terminology for intermediate standard[1]**	**Notes**	**Final selected quality standard**
Long-term	Water organisms	$QS_{fw, eco}$ $QS_{sw, eco}$	Refers to direct ecotoxicity	lowest water-based QS is selected as $AA\text{-}EQS_{fw}$ $AA\text{-}EQS_{sw}$
	Predators (secondary poisoning)*	$QS_{biota, secpois, fw}$ $QS_{biota, secpois, sw}$	QS expressed as concentration in biota	
		$QS_{fw, secpois}$ $QS_{sw, secpois}$	QS converted to corresponding concentration in water	
	Human health (consumption of fishery products)*	$QS_{biota, hh\ food}$	QS for water expressed as concentration in biota	
		$QS_{water, hh\ food}$	QS converted to corresponding concentration in water; valid for fresh and marine waters	
Short-term	Water organisms	$MAC\text{-}QS_{fw, eco}$ $MAC\text{-}QS_{sw, eco}$	Refers to direct ecotoxicity	$MAC\text{-}EQS_{fw}$ $MAC\text{-}EQS_{sw}$
D.w.	Human health (drinking water)		Relates to surface water used for abstraction of drinking water	$QS_{dw, hh}$

1: Note that the subscript ‘fw’ refers to freshwater, ‘sw’ to saltwater, ‘water’ is used for all waters, including brackish and marine waters.

1.5 National risk limits: SRC and NC

In addition to the WFD-standards, national specific environmental risk limits (ERLs) are derived for water. These are the Serious Risk Concentration (SRC) and the Negligible Concentration (NC) (see ERL Report 01).The NC is originally defined as the MPC/100. Given the similarity in protection aims, the AA-EQS/100 is used as the NC. One exception is made: if in the case of genotoxic carcinogens the final AA-EQS is based on human fish consumption, the AA-EQS will refer to an added cancer risk of 10^{-6} on a lifetime basis (see ERL Report 01, section 4.5). Since this is equal to the risk level that is originally associated with the Dutch NC, the additional factor of 100 will not be applied. In these cases, the protection level of the derived AA-EQS is already sufficient to meet the criteria of the NC.

1.6 Dutch standard characteristics

The derived ERLs will be expressed as dissolved concentrations in water. The methodology to recalculate water standards into a concentration in suspended matter makes use of the characteristics for Dutch standard suspended matter. For certain purposes, ERLs for water may have to be recalculated from dissolved to total concentrations. In that case, the same characteristics are needed. These characteristics are the percentage of organic carbon and the concentration of suspended matter in surface water. Guidance for recalculation methods is given in ERL Report 09.

1.7 Summary of input for QS-derivations

For an overview of the information that will be used for QS-derivation and triggers for the relevant routes to be considered, the following table format is advised. For information on A1-value and drinking water standards, see 3.4.

Table 2 Table format - collected properties relating to the assessments made in the report

Parameter	Value	Remark
log K_{oc}		needed for recalculation into total concentrations
BCF	xx L/kg	$QS_{water,\ secpois}$ relevant
BMF		$QS_{water,\ secpois}$ relevant
log K_{ow}		$QS_{water,\ secpois}$ relevant
H phrases[a]	H XX, H XX	$QS_{water,\ hh\ food}$ relevant
EU or WHO drinking water standard	XXX mg/L / not available	

a: Following implementation of Regulation (EC) 1272/2008 [7], the former R phrases have been replaced by H phrases. Since R phrases will be encountered regularly for substances, consult p. 20 of WFD guidance for translation of R to H phrases. For phrases not listed in WFD guidance, consult Annex VIII (p. 1352) of EC 1272/2007 (complete translation table).

1.8 Reader's guide

The following sections give additional guidance on derivation of the standards for direct ecotoxicity, secondary poisoning and human consumption of fish. If the latter two are triggered (see section 1.3.1 and 1.3.2), the lowest of the routes considered determines the final QS. In addition, a separate standard for freshwater used for drinking water abstraction. In principle, the derivation follows the WFD-guidance, but some aspects not (fully) covered are discussed here. Additional information is given on data presentation, evaluation and selection and the use of the collected information for standard derivation.

2 Derivation of WFD-standards for direct ecotoxicity

2.1 Collection and evaluation of aquatic laboratory toxicity data

For general guidance on collection and evaluation of ecotoxicity data, first consult ERL Report 02, Chapter 5. International guidelines exist for performing aquatic toxicity studies for many species. The most frequently used guidelines are summarised in Appendix 1.

2.1.1 Data tables for aquatic laboratory ecotoxicity studies

The ecotoxicity data are summarised in data tables. Separate tables are prepared for acute and chronic studies for freshwater and marine species. Marine species are defined as species living and tested in salt or brackish water. The division between freshwater, brackish water and seawater on basis of salinity is given in Table 3. The division in these categories is rather arbitrary and depends on the source used. For the division between freshwater and brackish or saltwater tests, the value of 0.5 ‰ is defined in the WFD [8].

Table 3 Classification of water according to salinity.

Water type	**Salinity (‰)**
freshwater	<0.5
brackish water	0.5 – 30
seawater	30 – 40

The following sections (2.1.1.1 to 2.1.1.18) discuss the parameters that are reported in the aquatic toxicity data tables, an example of which is presented in Table 4. The aim is to fill the table as complete as possible. The parameters are treated in the same order as they appear in the default toxicity data table. The following subsections have titles identical to the column titles in the data tables. Part of the text in this chapter is cited from Traas [9].

Table 4 Example of an aquatic acute ecotoxicity data-table for freshwater organisms.

Legend to column headings	
Species properties	relevant characteristics of the test species, such as age, size, origin
A	test water analysed Y(es)/N(o)
Test type	S = static; R = renewal; F = flow through; c = closed
Purity	refers to purity of active substance or content of active substance in formulation; ag = analytical grade; tg = technical grade
Test water	am = artificial medium; dtw = dechlorinated tap water; dw = deionised/dechlorinated/distilled water; nw = natural water; rw = reconstituted water; rtw = reconstituted tap water; tw = tap water
T	temperature
Ri	Reliability index according to [10]. Valid studies (Ri 2 or higher) are considered for EQS-derivation, depending on relevance and considering notes on data treatment (section 1.3.4)

Species	Species properties	A	Test type	Test compound	Purity [%]	Test water	Hardness $CaCO_3$ [mg/L]	pH	T [°C]	Exp. time	Criterion	Endpoint	Value [µg/L]	Ri	Note	Ref.
Bacteria																
Vibrio fischeri	strain NRRL-B-11,177	Y	S	active	ag				15	30 min	EC50	bioluminescence	61900	2	1	[a]
Cyanobacteria																
Anabaena flos-aquae		Y	S	product A	200 g/L					96 h	EC50		32800	4	2	[b]
Algae																
Desmodesmus subspicatus		Y	S	active	ag				21	72 h	EC50	growth rate	389000	2	3	[c]
Scenedesmus subspicatus	10000 cells/mL	N	S	active	tg			8.2-9.1	23	72 h	EC50	biomass	> 10000	3	4	[d]
Crustacea																
Asellus aquaticus	field collected	N		product A	200 g/L	am			10	1 h	NOEC	respiration	100	3	5	
Ceriodaphnia dubia	< 24 h	Y	S	product B	42.8%	dw	80-100		25	48 h	LC50	mortality	2.07	2	6	
Insecta																
Aedes aegypti	4th instar	N	S	active	97.4	dw			25	72 h	LC50	mortality	84	3	7	
Baetis rhodani	larvae, field collected	N	S	active	ag	am	180	7.4	15	48 h	LC50	mortality	8.49	3	8	
Amphibia																
Rana limnocharis	1 month old	N	R	active	> 95%	dw			20	96 h	LC50	mortality	82000	3	9	
Rana N. Hallowell	1.5 months old	N	R	active	> 95%	dw			20	96 h	LC50	mortality	129000	3	10	
Pisces																
Danio rerio		Y	S	active	ag	nw	140	8.4	21	96 h	LC50	mortality	241000	2	11	
Danio rerio		Y	S	product C	200 g/L	nw	140	8.4	21	96 h	LC50	mortality	214000	2	12	

Notes

1 Marine species, but tested in distilled water. Stability of test concentrations in distilled water checked for 21 d at 21ºC, room light. At 70 mg/L and lower stable for 21 d, at 105 and 140 mg/L stable for 17 d, thereafter decline to 84 and 76% of nominal, respectively.

2 Solvent 1% DMSO, solvent control included; no analysis of test concentrations, but short exposure time

3 etc.

2.1.1.1 Species

All available toxicity data for a given compound are ordered by test organism. Species are grouped in taxonomic groups. A comprehensive list of taxonomic groups is shown in ERL Report 11. Latin names are used for both taxa and species names. Species names within a taxon are listed in alphabetical order.

2.1.1.2 Species properties

The most relevant properties of the test organism are mentioned in this column; e.g. age, size, weight, life stage or larval stage. Toxicity data for organisms with different age, size, life stage etc., are presented as individual entries (i.e. one entry in each row) in the data table.

2.1.1.3 Analysed

This column reports whether the test compound is analysed during the experiment. Y (Yes) is entered in this column, when the compound has been analysed. When no analysis for the test compound is performed, N (No) is entered in this column.

In some cases the test compound is analysed (Y), but the test results (L(E)C50, EC10, NOEC) are not calculated from the actual concentrations. If the test result is based on nominal concentrations, this is mentioned in a footnote to this study: 'Test result based on nominal concentrations'. When this is valid because measured concentrations are close to initial concentrations (drop in concentration < 20% over exposure period), 'Test result based on nominal concentrations, measured concentrations were > 80% of nominal' is noted.

If the test results is based on nominal concentrations while there is considerable change in the concentration during the test (> 20% loss of test compound), the test result is recalculated using actual concentrations. In such case, in a footnote to this study should be mentioned that tests results were recalculated to actual concentrations. In static or renewal tests, when samples are analysed at different points of time, the mean of the measured values is used. When the initial concentration is not measured and one or more samples during the test are, a mean of the initial nominal and the measured concentration(s) is used. In general, taking the average of start and end concentrations slightly overestimates the average concentration during the whole experiment, while the geometric mean underestimates the concentration. For calculating the mean concentration during the course of a static experiment, the best assumption is an exponential decay of the concentration in time. In continuous flow experiments, the concentrations are usually reported as mean measured values, and here, no further calculations are necessary.

2.1.1.4 Test type

The following test types are distinguished:

- S static system
- Sc static system in closed bottles or test vessels
- R renewal system (semi-static)
- F flow-through system

CF continuous flow system
IF intermittent flow system

2.1.1.5 Test compound

This column can be deleted when the compound under consideration has only one structural molecular configuration.

If the tested compound is a metal, the tested metal salt should be reported here.

If the tested compound is a stereoisomer[2] , consists of a mixture of isomers, etc., the name of the tested molecule(s) should be reported here. For some stereo-isomers it might be preferred to derive individual risk limits. The stereoisomers dieldrin and endrin are an example of such a case.

If the tested compound is a structural isomer, the individual compounds, in general, have different physicochemical and toxicological properties and each compound will be subject of an ERL derivation (e.g. anthracene and phenanthrene).

Formulated products (e.g. biocides, pesticides) should be reported here.

Structural isomers
Compounds that are structural isomers are, in principle, regarded as different compounds, e.g. ethanol and dimethyl ether or anthracene and phenanthrene. In these cases, each individual isomer will generally be the subject of an ERL derivation. As a rule of thumb, isomers can be regarded as individual compounds when they have different CAS registry numbers. However, for more complex molecules[3] consultation with an expert or the client (e.g. the Ministry of IenM) might be needed.

2.1.1.6 Purity

Unit: %

The purity of the test compound expressed as percentage is reported in this column. Alternatively, the following abbreviations may be entered for the designation of chemical purity.

ag analytical grade
lg laboratory grade
pa pro analyse
rg reagent grade
tg technical grade

Here, the first four have a relatively high purity, while technical grade is in general somewhat less pure. When the purity of the test compound is expressed only by an abbreviation, this abbreviation is reported. However, a purity expressed as percentage is preferred.

[2] Stereoisomers: geometric isomers (cis- and trans-isomers or E- and Z-isomers), optical isomers (+ and – isomers or R- and S-isomers) and conformational isomers (e.g. chair and boat structures in cyclohexane ring structures)

[3] Isomers might be distinguished by CAS nrs., but still be treated (generally) as ‘one compound’, e.g. ‘nonylphenol’. The nonyl chain can have many conformations and different CAS nrs. exist. However, the generic name ‘nonylphenol’ is mostly used for all para-nonylphenol isomers.

2.1.1.7 Test water

In this column, the test water or medium is reported using abbreviations. Choose from the following list. A footnote to the test may be added if further description of the test medium is needed.

am: artificial medium, such as media used for bacterial and algal tests, artificial seawater
dtw: dechlorinated tap water
dw: de-ionised water, dechlorinated water or distilled water
nw: natural water, such as lake water, river water, sea water, well water
rw: reconstituted water: (natural) water with additional salts
rtw: reconstituted tap water: tap water with additional salts
tw: tap water

2.1.1.8 Hardness and salinity

Unit: mg $CaCO_3$/L

This column is shown in tables showing data from freshwater experiments, not for marine water. The hardness of the test water should be reported here. If the hardness of an artificial medium is not reported, but the composition of the medium is reported, the hardness should be calculated. Recalculation should be performed by summing the molar concentrations of all calcium (Ca) and magnesium (Mg) salts and expressing the result as $CaCO_3$ in units of mg/L.

2.1.1.9 Salinity

Unit: ‰

This column is only shown in tables showing data from saltwater experiments, and takes the place of the column for hardness in the freshwater tables. In practice salinity may be determined by recalculating the measured chloride ion only to total salinity, using the assumption that the total amount of all components in the oceans is constant. The average salinity of seawater is around 35 ‰ (roughly 35 g of salts per litre of seawater). The unit of salinity might also be found expressed in promille or ppt, i.e. parts per *thousand,* (ppt; not parts per trillion in this case) as w/w. To derive the salinity expressed in promilles the following conversion can be applied:

- when only chloride ions (Cl^-) have been measured, the salinity can be recalculated to ‰ from the chloride concentration using: S(ppt) = 1.80655 × chloride concentration (ppt), in which S = salinity;
- psu = practical salinity units4. One psu roughly equals one ppt (‰). Seawater has a salinity of approximately 35 psu ≈ 35 ‰ = 35 g/kg.

2.1.1.10 pH

If possible, measured pH values should be reported. If a pH range is given, this range is reported.

4 However, because of the qualitative nature in which salinity is used in EQS derivation, this definition and its inherent accuracy are not relevant.

2.1.1.11 Temperature

Unit: °C

In this column the temperature at which the test is performed should be reported, preferably a measured temperature. If a temperature range is given, the range is reported.

2.1.1.12 Exposure time

The duration of exposure to the toxicant in the toxicity experiment is expressed in this column. The abbreviations listed below in Table 5 can be used. A rule of thumb is to stick to the most common expression of test duration in case of standardised tests (e.g. OECD or ISO tests) where this is possible.

Table 5 Used abbreviation and applied range for exposure times.

Test duration in	**Abbreviation**	**Duration**
minutes	min	0-60 minutes
hours	h	1-120 hours
days	d	5-56 days
weeks	w	1-4 weeks
months	mo	1-12 months
years	y	≥ 1 years

2.1.1.13 Criterion

This column gives the effect measure that is reported, e.g. NOEC, EC50, EC10. For interpretation and use of these test results, see ERL Report 02, Section 5.2.2.

2.1.1.14 Test Endpoint

This column reports the test endpoint studied, e.g. mortality, growth, immobilisation. More information on relevant endpoints is given in ERL Report 02, Section 5.2.3.

2.1.1.15 Value

Unit: mg/L, µg/L.

The unit in which the results of toxicity tests are expressed is optional. For reasons of comparison and to avoid errors, the same unit is used throughout all aquatic toxicity data tables in one report. In general, values are expressed in the number of digits reported in the study, usually two or three digits. At most, four significant digits are reported. However, further calculation with these data may be necessary: averaging, dividing the values by an assessment factor, use of the results in SSDs, etc. Further calculation is always performed with the original (not rounded off) values.

Toxicity data of metal compounds are always expressed in quantities of the element, not as the salt. For example, a test performed with $CoSO_4.7H_2O$ is expressed as Co_2^+. Test results are recalculated if necessary. A similar approach is followed for all charged substances with a non-toxic counter ion.

2.1.1.16 Reliability

This column contains a number (1, 2, 3 or 4), indicating the quality of the study summarised according to section 2.2.

2.1.1.17 Notes

This column contains references to footnotes that are listed below the toxicity data tables. Numbers are used to refer to footnotes.

2.1.1.18 Reference

The reference to the study from which data are tabulated, All cited references are listed in a reference list. If references are generated using bibliographic software (e.g. Endnote, Procite), it is most convenient to list all references, including those of the Annexes, into one single reference list.

2.2 Selection and aggregation of aquatic laboratory data

Where multiple data are available for the same species/endpoint that are obtained under comparable test conditions, individual toxicity data may be aggregated using the same principles as those in Chapter R.10 of the REACH Guidance [11]. This aspect is discussed in general terms in ERL Report 02, Section 5.3 and are supplemented here with specific guidance for aquatic data. This process is performed separately with toxicity data for freshwater species and marine species (see also 2.3).

For non-standard test species, preference is given to endpoints that are applicable to related standard test species, such as emergence, growth and survival or biomass. If for a species only alternative endpoints are available, these may be used, although this has to be judged on a case-by-case basis.

If endpoints are available for different durations, preference is given to the endpoints from tests that followed the minimum test duration as specified in the guideline, e.g. at least 72 hours for algae, 48 hours for daphnids, 96 hours for fish. If for *D. magna* endpoints are available from 24- and 48-hours test, the latter is preferred for risk assessment even when it is higher than the 24-hours value, since a test duration of 48 hours is prescribed in the guideline. In principle, the test duration for daphnids is considered applicable to other invertebrates as well.

If there is a clear relationship between test results and abiotic conditions, results are selected that refer to conditions relevant for Dutch surface waters. Any deselection of data should be motivated. The aggregated data should be presented in a new table, according to the format shown below. The selected acute and chronic values are presented separately for each species, and a footnote is added to explain how the value is derived from the summary data tables.

Table 6 Example of an aggregated data-table with selected acute and chronic ecotoxicity data for freshwater organisms.

Acute			**Chronic**		
Taxon/species	**L(E)C50 [µg/L]**	**Ref.**	**Taxon/species**	**NOEC/L(E)C10 [µg/L]**	**Ref.**
Bacteria			**Algae**		
Vibrio fischerii	58876 [a]	[a]	*Desmodesmus subspicatus*	106000 [k]	
V. qinghaiensis sp.	79255	[b]	*Pseudokirchneriella subcapitata*	<100000 [c]	
Algae			**Crustaceans**		
Desmodesmus subspicatus	389000 [b]	[c]	*Asellus aquaticus*	1.35 [d]	
Pseudokirchneriella subcapitata	>100000 [c]	[d]	*Daphnia magna*	1768 [l]	
Crustaceans			*Gammarus pulex*	2.95 [d]	
Asellus aquaticus	119 [d]	[f]	*Hyallella azteca*	0.47 [h, m]	
Ceriodaphnia dubia	2.07	[g]	**Insects**		
Chydorus sphaericus	832	[f]	*Caenis horaria*	0.024 [d]	
Cypretta seuratti	1	[h]	*Chaoborus obscuripes*	1.99 [d, m]	
Cypridopsis vidua	10 [d]	[f]	*Chironomus riparius*	< 0.4 [c,n]	
Daphnia magna	52455 [e]	[i,j]	*Chironomus tentans*	0.42 [m]	
Gammarus pulex	110 [d]		**Fish**		
Insects			*Danio rerio*	300000	
Caenis horaria	1.77 [d]		*Oncorhynchus mykiss*	1200 [p]	
Chaoborus obscuripes	284 [d]				
Chironomus dilutus	2.65				
Chironomus tentans	6.9 [g]				
Cloeon dipterum	1.02 [d]				
Epeorus longimanus	0.65 [h]				
Fish					
Danio rerio	227099 [j]				
Leuciscus idus melanotus	237000				
Oncorhynchus mykiss	211000				
Annelids					
Lumbriculus variegatus	6.2				

a: geometric mean of 61900 and 56000 for tests with active and formulation; marine species tested in freshwater
b: test with active, endpoint for formulation >3 times lower
c: unbound values are not used for EQS-derivation, value included to show that species has been tested
d: lowest relevant endpoint, immobility
e: geometric mean of 30000, 85000, and 56600, 48 h tests with formulation and active, endpoint immobility
etc.

2.3 Combining freshwater and marine data sets for ERL derivation

After compiling the aggregated data table, it should be investigated whether toxicity data for freshwater and for marine species may be combined into one (aggregated) data table. Data on organic substances are pooled, unless there are reasons not to do so. E.g. differences in osmoregulation may cause differences in toxicity for organic compounds [12], however the decision is taken based on the available toxicity data. If a different sensitivity is not apparent from the data collected, data are pooled for ERL derivation. For metals, data are kept separated a priori because it is likely that differences in metal speciation

and osmoregulation will cause differences in toxicity. However, pooling may be allowed if a difference cannot be demonstrated. Extensive guidance on this point is given in the WFD-guidance (see below for relevant sections).

If fresh- and saltwater data are pooled, the standards for both freshwater and marine water are derived using the same, combined dataset, but with different assessment factor schemes for the AF- and SSD-approach. By default, an additional assessment factor of 10 is applied for the marine assessment as compared to freshwater assessment. This additional assessment factor can be decreased in a stepwise manner when toxicity data for specific marine species or taxa are available. An additional factor of 5 is used if the dataset contains one typically marine species. The WFD-guidance specifies how this should be interpreted. If two or more specifically marine species are present, the freshwater and marine assessment schemes are similar. Note that this does only apply to the AF- and SSD-approach, and not to the mesocosm approach (see 2.6.4). When the freshwater and marine data cannot be pooled for QS derivation, the separate aggregated data sets are used for QS-derivation.

➔ Location in WFD guidance: Section 3.2.3, p. 35.
➔ Location in WFD guidance: Section 3.3.2.1, p. 46.
➔ Location in WFD guidance: Appendix A1.3.7.1, p. 151.

2.4 Additional guidance on the Assessment Factor approach

2.4.1 Presence of sensitive taxa

The quantity and type of data available determines the assessment factors used. The assessment schemes for derivation of the QS_{eco} and MAC_{eco} are presented in detail in the WFD-guidance [1]. The schemes have been developed for all types of chemicals, including those for which ecotoxicity data are scarce, and offer the possibility to derive a QS and MAC in case only acute data for algae, *Daphnia* and fish are available. The use of an AF of 10 on the lowest NOEC or EC10 is allowed if additional chronic NOECs (EC10) are present for three species from three trophic levels, provided that the species tested represent one of the more sensitive taxonomic groups. This is made clear in one of the footnotes to the table with assessment factors for freshwater in the WFD-guidance, which states (footnote d):

'…When examining the results of long-term toxicity studies, the $QS_{fw, eco}$ should be calculated from the lowest available long term result. Extrapolation to the ecosystem can be made with much greater confidence, and thus a reduction of the assessment factor to 10 is possible. This is only sufficient, however, *if the species tested can be considered to represent one of the more sensitive groups*. This would normally only be possible to determine if data were available on at least three species across three trophic levels.'

The link with trophic level that is made in the assessment schemes is complicating. Crustaceans and insects may belong to the same trophic level, while depending on the mode-of-action large differences in sensitivity may exist between these taxonomic groups. Similarly, the

primary producers algae and macrophytes may also show large differences in sensitivity. For the choice of the assessment factor, the availability of data for the potentially most sensitive taxonomic group is most important, rather than having three trophic levels [2].

Determining whether or not the potentially sensitive species group is included in the dataset may be difficult [13]. Given the fact that test results for the same species may easily differ by a factor of 10, the question is which difference between test results should be considered as indicative for a taxon-related difference in sensitivity. As a pragmatic approach, if the lowest test results per taxon differ by more than a factor of 3, this is considered as an indication that one taxon is more sensitive than another. If this sensitive taxon is not represented in the chronic dataset, a higher assessment factor should be applied. All additional relevant information that substantiates the choice of the assessment factor should be considered, including information from additional (non-standard) studies, read-across and QSAR-data [2]. Information from e.g. mesocosm studies may also point at sensitive taxa that are not adequately represented in the laboratory dataset. This may lead to a higher assessment factor than originally selected on the basis of the schemes alone. Guidance on the differentiation between taxonomic groups is given in ERL Report 10.

2.4.2 Use of endpoints for micro-organisms

According to the WFD-guidance [1], data for bacteria representing a further taxonomic group may only be used if non-adapted pure cultures were tested. Furthermore, studies with bacteria (e.g. growth tests) are regarded as short-term tests. Consequently, the WFD-guidance states that unlike for algae, NOECs or EC10-values derived from bacterial studies may not be used in the derivation of the AA-EQS using assessment factors. EC50 values from bacterial tests may be used as additional acute data. If, however, a reliable bacteria test is available that is comparable to an algae test in terms of duration and endpoint (i.e. 72 hours and specific growth rate), there is no scientific reason to exclude such endpoints from the dataset. The same principle applies to toxicity data using protozoans. For the purpose of national EQS-derivation, NOECs and EC10-values for bacteria and protozoans are accepted as chronic endpoints, if obtained in a comparable way as those for algae.

The WFD-guidance does not make reference to fungi as a specific taxonomic group. Data on fungi are considered relevant for fungicide risk assessments and may become available in the (near) future. If growth tests with fungi are present, it is advised for the time being to treat the data similarly to algae, i.e. include the EC50 for the acute dataset and the NOEC in the chronic dataset.

2.5 Additional guidance on SSDs

2.5.1 Data requirements

The WFD-guidance gives criteria for construction of a Species Sensitivity Distribution (SSD), which are in accordance with REACH guidance [11]. According to the guidance, the output from an SSD-based quality standard is considered reliable if the database contains preferably more

than 15, but at least 10 data points, from different species covering at least the following eight taxonomic groups:

- Fish
- A second family in the phylum Chordata
- A crustacean
- An insect
- A family in a phylum other than Arthropoda or Chordata
- A family in any order of insect or any phylum not already represented
- Algae
- Higher plants

If freshwater and marine datasets cannot be pooled, the requirements should be met for each of the two datasets. However, some of the taxa mentioned above (e.g. insects, higher plants) are not (well) represented in marine environments and may be replaced by other taxa. Also in that case, the minimum number of taxa and data points should be met [1]. In some cases, where a large dataset is available but one of the listed taxa is missing, it may be considered to use SSDs. In this case, using only the lowest endpoint with an assessment factor would mean that a lot of valuable information is neglected. For plant protection products with a specific mode of action, additional data will most often focus on the potentially sensitive species groups. For insecticides, authorisation dossiers will most often not contain data on aquatic macrophytes, since only data for algae are required. Similarly, data on insects may not be included in the dataset for herbicides, since only crustaceans are required for authorisation. Moreover, because reduction of vertebrate testing is an important issue, authorisation dossiers may no longer include multiple fish studies in the future. Examples of justification of the use of SSDs for datasets that did not fully meet the requirements can be found in several RIVM-reports [14-18].

2.5.2 Constructing specific SSDs

The WFD-guidance offers the option to derive quality standards on the basis of specific SSDs for sensitive taxonomic groups. In this case, the minimum number of data points (10, preferably 15) should be maintained. However, this is only possible if from a generic SSD with all required taxa there is clear evidence of a ‘break’ in the distribution between the sensitive and other species (bimodal distribution), or if there is poor model fit [1]. In principle, this guidance is followed, meaning that the mode of action alone is not enough reason to justify an SSDs for a potentially sensitive group, without having data on the above listed required taxa. In other words, the minimum requirements to perform an SSD should also be met for a compound with a specific mode of action, in order to be able to demonstrate deviations from the expected distribution [2].

However, it is recognised that for herbicides and insecticides often large datasets are present for sensitive species groups, while data for other taxa are missing. Sometimes, it may be possible to demonstrate a break in the distribution for the acute dataset, but too many taxa are missing to construct a generic chronic SSD. In this case, it may be considered to apply specific SSDs for both datasets. However, the implicit assumption

that the sensitivity distribution on the acute and long term time scale is similar has not been proven yet. Comparing the position of specific taxa in species sensitivity distributions between acute and chronic SSDs is an important topic for future research [2].

In principle, a specific SSD is made for the most sensitive taxon, but multiple sensitive taxa may be combined in one specific SSD based on the distribution of data. This specifically applies to insecticides for which insects and crustaceans may be combined in one SSD for arthropods if the data show that sensitivities of the respective species groups overlap [2]. If a specific SSD is constructed, it should always be checked if the result is sufficiently protective for taxa that were not included in the SSD.

2.5.3 Assessment factors for a specific SSD

For derivation of the $QS_{fw, eco}$, a default assessment factor of 5 is applied to the Hazardous Concentration for 5% of the species (HC5) that is derived from an SSD based on chronic ecotoxicity data. The WFD-guidance lists five topics that are relevant when considering a lower factor:

- the overall quality of the database and the endpoints covered, e.g., if all the data are generated from “true” chronic studies (e.g., covering all sensitive life stages);
- the diversity and representativeness of the taxonomic groups covered by the database, and the extent to which differences in the life forms, feeding strategies and trophic levels of the organisms are represented;
- knowledge on presumed mode of action of the chemical (covering also long-term exposure). Details on justification could be referenced from structurally similar substances with established mode of action;
- statistical uncertainties around the HC5 estimate, e.g., reflected in the goodness of fit or the size of confidence interval around the 5^{th} percentile, and consideration of different levels of confidence (e.g. by a comparison between the median estimate of the HC5 with the lower estimate (90% confidence interval) of the HC5);
- comparisons between field and mesocosm studies, where available, and the HC5 and mesocosm/field studies to evaluate the level of agreement between laboratory and field evidence.

Based on case studies, some examples for justifying a lower factor are given in [2]. A default factor of 10 is used for the SSD-based MAC-$QS_{fw, eco}$. This factor may be adapted if other lines of evidence suggest that a higher or lower one is appropriate. Such evidence may consist of information on the ratio between acute L(E)50 and EC10/NOEC-values and the topics that are listed in the section on the $QS_{fw, eco}$. When specific SSDs are constructed for sensitive species groups, some of the uncertainty described in the WFD-guidance still remains and should be addressed, however, lowering the assessment factors is reasonable because uncertainty about the representativeness of the tested species is reduced.

The scheme in Table 3 is proposed in [2] and may be used as a starting point for derivation of standards for freshwater. If a pooled dataset is used, the corresponding saltwater standards are derived using an additional assessment factor of 10, which can be decreased to 5 if one typically marine species is represented in the dataset. If at least two typically marine species are present, no additional assessment factor is needed for the saltwater assessment.
An important note is that when deriving an SSD-based MAC-$QS_{fw, eco}$ using L(E)C50-values, an assessment factor >1 is needed because the SSD-result relates to a 50% effect level, whereas the MAC-$QS_{fw, eco}$ refers to no effects.

Table 7 Assessment factors to be put on a HC5 to derive freshwater standards based on different types of datasets [2]. Shaded cells represent the values as given in the WFD-guidance [1]. Saltwater standards are derived using an additional AF of 10 or 5, depending on the presence of typically marine species.

	$QS_{fw, eco}$	MAC-$QS_{fw, eco}$	
	input: chronic NOEC/EC_{10}	input: acute L(E)C_{50}	input: acute NOEC/L(E)C_{10}
generic SSD	**default 5** **range 5-1**	**default 10**	default 5 range 5-1
specific SSD	default 3 range 3-1	default 6 range 6-2	default 3 range 3-1

2.6 Using mesocosm data for QS-derivation

➔ Location in WFD guidance: section 3.3.1.3, p. 43-45.

2.6.1 Assessment of reliability

Aquatic micro- and mesocosm studies are frequently submitted in the context of registration of agricultural pesticides. According to the WFD-guidance, they are a valuable tool to assess the impact of a chemical on populations or communities of aquatic ecosystems under more realistic environmental conditions than is achievable with standard single species laboratory studies. If reliable mesocosm data are available, they may be used either as the basis of aquatic ERLs, or used as additional information for the selection of the assessment factor applied to an SSD [1]. General guidance on the design of mesocosm studies is given in several documents [19-22]. A guidance document on the evaluation and interpretation of study results was published in 2008 [23].
According to this guidance, the following questions should be answered to assess the reliability of mesocosm studies:

- Is the test system adequate and does the test system represent a realistic freshwater community?
- Is the description of the experimental set-up adequate and unambiguous?
- Is the exposure regime adequately described?
- Are the investigated endpoints sensitive and in accordance with the working mechanisms of the compound, and with the results of the first-tier studies?
- Is it possible to evaluate the observed effects statistically and ecologically?

To facilitate answering these questions, the guidance provides more information on the aspects to be considered and contains a detailed checklist to assess the scientific reliability of the study [23]. A critical part of the evaluation of mesocosm studies is the statistical analysis of measurement endpoints related to effects. Various univariate and multivariate techniques are available for evaluation of effects at the population and at the community level. Detailed information on methodology and statistical evaluation can be found in [23] and references therein.

In 2013, the European Food Safety Authority (EFSA) published guidance on the risk assessment of plant protection products (PPP) [24]. This guidance elaborates on the aforementioned guidance documents and specifically addresses the set-up, evaluation and use of mesocosm studies for risk assessment of PPP in edge-of-field surface waters. For example, EFSA specifies that, besides representatives of different trophic levels, at least 8 different populations of the sensitive taxonomic group need to be present in the micro-/mesocosm test systems and for which a concentration–response relationship can be derived.

Regarding statistical evaluation, detailed information on EFSA introduces the Minimum Detectable Difference (MDD) as an additional criterion. The MDD expresses the difference between control and treatment that can be detected as significant, given a specific test design and control performance. The MDD is particularly important if no effect is observed, since when a LOEC can be calculated the statistical power apparently is high enough to detect an effect. However, if the MDD is >100%, due to e.g. low abundance or variability in the control, it is not possible to derive a meaningful NOEC, since in this case it is not possible to underpin statistically that there is no difference between treatment and control [24]. EFSA advises that the MDD is reported for each measurement endpoint and states that the MDD should preferably be lower than 70-90%. It is noted that for field studies with earthworms and non-target arthropods, a lower level of 50% effect should be detectable [25,26]. However, EFSA [24] also requires that for at least 8 sensitive taxa a statistical evaluation of the dose-response relationship should be possible, meaning that the MDD should be sufficiently low. The case study with an insecticide that is included in the EFSA guidance shows that low MDDs for sensitive endpoints are indeed possible.

2.6.2 Effect classes

If a study based on the abovementioned criteria is considered reliable, Effect classes are used to summarise the observed effects in a transparent and comparable way. The original classes were developed by [27,28] and adapted later on [23,24]. The Effect classes are summarised as follows (after EFSA):

Table 8 Description of Effect classes used to classify effects in mesocosms.

Effect class	Description
0	*Treatment related effects cannot be evaluated*. Due e.g. low abundance and variability the MDD was always larger than 100 % so even very strong effects could not be determined for the endpoint evaluated. If this class is consistently assigned to endpoints that are deemed most relevant for the interpretation of the study, the regulatory reliability of the micro-/mesocosm tests is questionable.
1	*No treatment-related effects demonstrated for the most sensitive endpoints*. No (statistically and/or ecologically significant) effects observed as a result of the treatment. Observed differences between treatment and controls show no clear causal relationship.
2	*Slight effects* Effects concern short-term and quantitatively restricted responses usually observed at individual samplings only.
3A	*Pronounced short-term effects (< 8 weeks, followed by recovery)* Clear response of endpoint, but full recovery of affected endpoint within 8 weeks after the first application or, in the case of delayed responses and repeated applications, the duration of the effect period is less than 8 weeks and followed by full recovery[5]. Treatment-related effects demonstrated on consecutive samplings.
3B	*Pronounced effects and recovery within 8 weeks post last application* Clear response of the endpoint in micro-/mesocosm experiment repeatedly treated with the test substance and that lasts longer than eight weeks (responses already start in treatment period), but full recovery[5] of affected endpoint within eight weeks post last application.
4	*Pronounced effect in short-term study* Clear effects (e.g. large reductions in densities of the population) observed, but the study is too short to demonstrate complete recovery within eight weeks after the (last) application.
5A	*Pronounced long-term effect followed by recovery* Clear response of sensitive endpoint, effect period longer than 8 weeks and recovery did not yet occur within 8 weeks after the last application but full recovery[5] is demonstrated to occur in the year of application.
5B	*Pronounced long-term effects without recovery* Clear response of sensitive endpoints (> 8 weeks post last application) and full recovery cannot be demonstrated before termination of the experiment or before the start of the winter period.

[5] An endpoint is considered as recovered if the MDD allows statistical evaluation during the relevant recovery period (so excluding MDD class 0) and the conclusion of no statically significant effect between treated systems and controls is not caused by a decline of that endpoint in controls (e.g. at the end of the growing season). If these criteria are violated a higher effect class has to be selected.

The Effect classes are assigned to all different endpoints measured in the study, e.g. abundance of specific taxa based on univariate statistics, diversity indices or community endpoints based on multivariate analyses. A summary of the Effect classes is made to enable the overall assignment of Effect classes to the respective treatments (Figure 1).

Box 2 Example of the summary of the Effect classes observed for several endpoints in the outdoor microcosm study with xxxx; ↓ indicates a downward trend; ↑ indicates an upward trend. *TWA*, Time weighted average, *PRC*, principal response curve

	Water concentration (µg a.s./L)				
Nominal concentration:	3	15	30	150	300
Measured peak concentration:	2.8	14.5	28	146	292
7-day TWA concentration:	2.5	12.9	24.9	130	260
21-day TWA concentration:	2.0	11.5	22.2	116	231
Species/group					
Chlorophyll a – periphyton	1	1	1	1	1↑
Chlorophyll a – phytoplankton	1	1	1	1	1↑
Periphyton (PRC)	1	1	1	1	1↑
Periphyton (populations)	1	1	1	1	2↑
Phytoplankton (PRC)	1	1	3A↓	3A↓	3A↓
Phytoplankton (populations)	1	1	3A↓	3A↑↓	3A↑↓
Zooplankton (PRC)	1	1	3A↓	5A↓	5B↓
Zooplankton (populations)	1	2↓	3A↓	5A↓	5B↓
Macroinvertebrate, sweep net (PRC)	1	1	1	1	4↓
Macroinvertebrate populations	1	1	1	1	5B↓

Figure 1 Illustration of a summary of Effect classes for various endpoints measured in five mesocosm treatments. Based on this overview, Effect class 1 is assigned to the nominal concentration of 3 µg/L, Effect class 2 to 15 µg/L. Box copied from [23]

2.6.3 *Use of Effect classes for EQS-derivation*

The WFD-guidance only refers to the NOEC of a mesocosm, but does not make reference to the Effect classes. It is stated, though, that ecological recovery is not considered when deriving aquatic EQSs within the context of the WFD (see WFD-guidance [1], section 2.8.2, 3.3.1.3). In a Dutch proposal for aquatic effects assessment of pesticides [2] additional guidance is given on the use of mesocosm data for EQS-derivation. According to this guidance at least Effect Class 3 concentrations and higher are not relevant for EQS-derivation, because an initial treatment-related effect on a relevant ecological endpoint is demonstrated. Strictly speaking, Effect Class 1 concentrations are equal to the NOEC, since at that concentration no consistent and statistically significant treatment-related effects are found. According to [2] Effect class 2 concentrations may be used as well, since they relate to situations in which 'treatment-related effects are reported as 'slight', 'transient', or other similar descriptions. It concerns a short-term and/or quantitatively restricted response of one or a few sensitive endpoints, usually observed at individual samplings only.' Application of a larger assessment factor to Effect Class 2 concentrations may ensure appropriate protection and a cost-effective use of micro-/mesocosm experiments [2].

2.6.4 Treatment of freshwater and saltwater data

Little information is present on the representativeness of freshwater studies for marine risk assessments. Differences in physico-chemical characteristics, water exchange rate and sensitive taxa may contribute to differences in ecological response. According to the WFD-guidance (section 3.3.2.3), freshwater mesocosm studies may be used as a basis for a marine risk assessment, but an additional assessment factor of 10 should be applied in line with the AF-approach (see 2.4). Supplementary to the WFD-guidance, it may be considered to lower the additional assessment factor if the laboratory dataset indicates that the sensitivity of the typically marine species is covered by the freshwater species (i.e. the effect levels for the typically marine species are in between those for freshwater species).

Regarding the use of marine mesocosms for the derivation of a freshwater EQS, it should be noted that according to the WFD-guidance marine mesocosm data often apply solely to small pelagic organisms. It should be considered that such studies may therefore seriously under-represent many taxa, e.g. benthic epifauna and macrophytes. On the other hand, marine mesocosms may point at sensitive taxa that are not represented in the freshwater dataset (e.g. molluscs). If for the laboratory dataset it is decided that freshwater and marine data can be pooled, there is no scientific objection to use a valid marine mesocosm also in the freshwater assessment. However, if the critical endpoint in the marine mesocosm is for a typically marine taxon which has no freshwater representatives (e.g. Echinoderms), the representativeness of the result for a freshwater assessment should be carefully considered, e.g. considering the size and diversity of the freshwater dataset in relation to the diversity in the mesocosm. Brock et al. [2] advice that a single marine mesocosm should not be used as the sole basis for a freshwater standard.

It should be recognised that when the freshwater data set is small, e.g. three taxa are represented, the addition of a lower toxicity value to the data set when only one new toxicity value is added is already 25% by chance alone. This means that the new, lower value may be found for a freshwater or a saltwater species, or any new taxon, by equal probability, assuming that we have no specific information on the mode of action that allows to pinpoint the sensitive species a priori. Hence, for small (but pooled) data sets, if the critical endpoint is derived for a typically marine taxon (no freshwater representative) the standard for the freshwater data set will be derived also on a data set that includes this specific marine organism.

However, if the freshwater toxicity data set is extensive (e.g. it was possible to construct an SSD) and ample information on the sensitivity of various taxa is available, but a critical toxicity value for a typical marine species is derived that is significantly lower than all freshwater data, it could be considered to leave the toxicity data for the exclusively marine species (NB that has no freshwater equivalent) out of the freshwater data set.

The following sections discuss the use of mesocosm results for derivation of the QS or MAC-QS, which is considered applicable to freshwater and saltwater mesocosms. However, for the ease of reading, only the subscript for freshwater is used.

2.6.5 Assessment of exposure

The evaluation and selection of mesocosm data as discussed in the previous sections results in identification of Effect Class 1 and/or 2 treatments that may be used for derivation of the $QS_{fw, eco}$ or MAC-$QS_{fw, eco}$ and respective Whether or not a particular mesocosm study is indeed relevant depends on the exposure regime that was applied in the study. Basically, the same considerations have to be made as for laboratory tests: the exposure conditions should match the purpose of QS-derivation, e.g. the $QS_{fw, eco}$ should preferably be based on studies with long-term continuous exposure, whereas studies with peak exposure may be used for derivation of the MAC-$QS_{fw, eco}$. However, since existing mesocosms for pesticides have often been submitted for authorisation of PPP, they are designed to reflect the agricultural use and may not (fully) meet the requirements for QS-derivation. Studies may simulate single or replicated applications and depending on the dissipation rate, the following exposure patterns may be found in the water phase:

- single pulse with decline of concentrations to 0 within a few days
- single pulse with decline to 0 within days to weeks
- multiple pulses with decline to 0 in between applications
- multiple pulses with accumulation of concentrations between applications

Guidance on how these patterns may be used for EQS-derivation is based on [2,3,24]. There are two issues:

- the duration of exposure should reflect the relevant duration in the field, i.e. a short-term peak for the MAC-$QS_{fw, eco}$ and long-term exposure for the $QS_{fw, eco}$
- the concentration in the mesocosm that is associated with the no-effect level should be adequately expressed, i.e. a choice should be made between nominal, measured peak of time weighted average (TWA) concentrations

Regarding the relevant duration, it is advised that studies involving single or multiple pulses with a relatively fast decline can only be used for derivation of the MAC-$QS_{fw, eco}$. For the $QS_{fw, eco}$, the substance should have been present in the water phase for a longer period of time. For tests with multiple applications of fast dissipating substances, it is stated in [2] that concentrations should not drop below 10% of the peak concentration in between applications, while tests with single pulses can only be used when dissipation rate is relatively slow, but no further guidance on dissipation rate is given. Following EFSA [24], a single pulse study can only be used for chronic QS-derivation when the concentration has not declined to levels lower than 20% of nominal during the time-window that is used for calculating the TWA concentration that is associated with the NOEC-treatment (Effect class 1 or 2, see 2.6.3). This relates to the second issue, the expression of the results.

For QS-derivation it is advised to express the no-effect treatment in a mesocosm on the basis of a TWA concentration. The length of the TWA time window should be guided by the length of the relevant critical test from the laboratory dataset, i.e. the test that delivered the lowest L(E)C50 or NOEC/EC10. Additional information on the time to onset of maximum effects, the length of the most sensitive life stage, the acute to chronic ratio may be used to further underpin or adapt the choice of the time window. For example, for derivation of the $QS_{fw, eco}$, the mesocosm-NOEC is initially derived by expressing the Effect class 1 or 2-treatment on the basis of 21-days TWA if the 21-days NOEC for *Daphnia magna* was the lowest endpoint from the laboratory dataset. However, if in the treatment level above the level identified as Effect Class 1 the time to onset of maximum effects is 30 days, the NOEC of the Effect class 1 treatment should be calculated as the 30-days TWA concentration. For derivation of the MAC-$QS_{fw, eco}$, the Effect class 1 or 2-treatment is expressed on the basis of the 48- or 72-hours TWA after the highest peak, depending on whether arthropods or algae are most sensitive in the laboratory tests.

Taking this criterion as a starting point, this means that if the lowest laboratory NOEC is obtained from a 21-days *Daphnia* study, a single pulse mesocosm can only be used for a $QS_{fw, eco}$ if the concentration in the water phase during 21 days is at least 20% of the initial peak. In this case, the DT50 for dissipation from the water phase in the mesocosm should have been 9 days or higher. Figure 2 gives a graphical representation of an Effect class 1 treatment not meeting the criterion (left hand side) and one just meeting this criterion (right hand side).

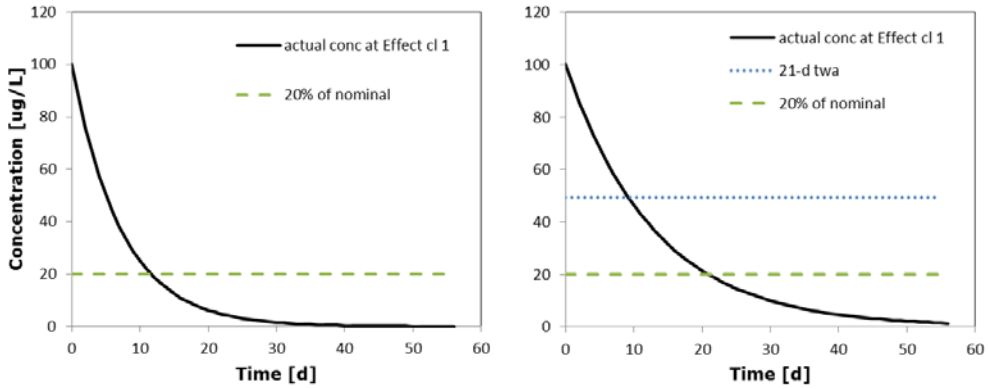

Figure 2 Development of concentrations in a single pulse Effect class 1 mesocosm-treatment. The critical laboratory test is a 21-d Daphnia test. The initial concentration is 100 µg/L. The green dashed line represents 20% of initial. The treatment at the left hand side does not meet the criterion, because after 21 days, the concentration has declined to 5% of initial.The treatment at the right hand side meets the criterion of 20% of initial left after the critical time window and the NOEC is expressed as the 21-days TWA (blue dotted line).

If decline is faster than required, studies may still be used for derivation of the $QS_{fw, eco}$, provided that repeated dosing is applied and the concentrations in between treatments does not fall below 20%. In addition, the application period should be long enough to cover the required time window (Figure 3). The appropriate effect class

concentration is then calculated as the time weighted average concentration over the test period.

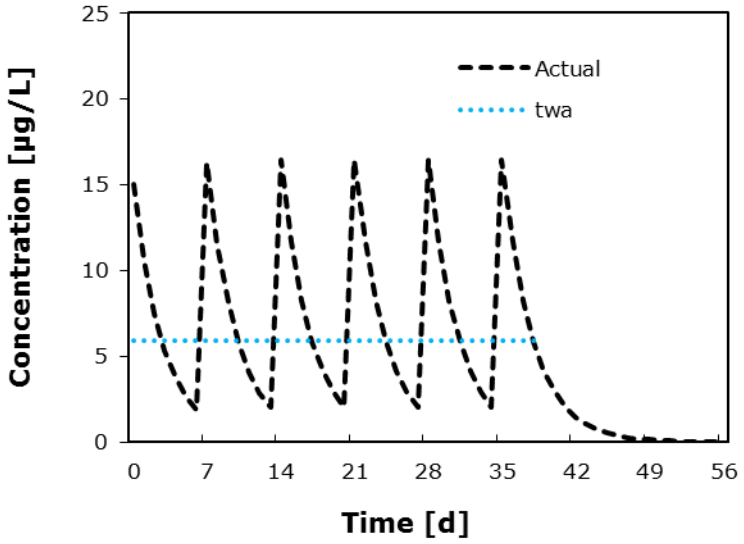

Figure 3 Development of concentrations after repeated applications, minimum concentration between dosing is >20% of initial and the application period is longer than the critical laboratory test of 21 days.

The requirement of at least 20% of initial left within the appropriate time window can also be applied to the MAC-$QS_{fw, eco}$: with a time-window of 48 to 72 hours, the minimum DT50 for dissipation from the water phase should be 0.9 to 1.3 days, respectively. Studies with multiple applications can be considered as a worst case exposure regime for derivation of the MAC-EQS, which may be reflected in the choice of the assessment factor [2].

If the concentration of a substance has fallen below 20% of initial within the appropriate time window, a case-by-case decision has to be made, by e.g. considering the time to effect in the laboratory or mesocosm studies. If a shorter time window is not appropriate, the test cannot be used for EQS-derivation, unless repeated dosing is applied.

2.6.6 *Assessment factors to be used on mesocosm results*

The WFD-guidance gives a default assessment factor of 5. A more differentiated assessment factor scheme is given in Table 9, based on [2]. The height of the assessment factor is always based on expert judgement considering all available information.

Table 9 Assessment factors for mesocosm studies [2].

TWA concentration associated with	**$QS_{fw, eco}$**	**MAC-QS**
NOEC = Effect class 1 for most sensitive structural endpoint	2-4*	1-2* (multiple applications) 2-3* (single application)
Effect class 2 for most sensitive structural endpoint	4-5*	2-3* (multiple applications) 3-4* (single application)

If a single adequate study is available, the higher assessment factor is used. If several adequate micro/mesocosm studies are available the assessment factor is applied to the highest test result or the lower

assessment factor is applied to the most critical test result. Since mesocosms generally do not contain fish, it should always be checked if the resulting mesocosm-based QS is protective for fish.
If a mesocosm study cannot be used as such for derivation of the QS or MAC, it may still be useful to underpin the choice of the assessment factor for the AF- or SSD-method.

As indicated above (see section 2.6.4), an additional assessment factor may be needed when using a freshwater mesocosm for the derivation of the $QS_{sw, eco}$.

2.7 Selection of the QS for direct ecotoxicity

When the SSD- and/or mesocosm approach is used next to the AF-approach, the $QS_{fw, eco}$, $QS_{sw, eco}$, MAC-EQS_{fw} and MAC-EQS_{sw} should be selected based on a comparison of the resulting values. This should include a discussion on the residual uncertainty, an explanation for the difference between the respective derivation methods and the reason for choosing the final method.

According to the WFD-guidance the final value should preferably be based on the results from the SSD- or mesocosm-approach, since these entail a more robust approach towards assessing ecosystem effects. The SSD gives an estimate of the range of sensitivities to be encountered in an ecosystem, but it is still based on single species laboratory data, and species-interactions at the ecosystem level that may occur in the field n the field are not included. Mesocosm studies most often do not allow for determination of exact cause-effect relationships, but may point to long-term ecosystem effects that cannot be shown in single-species laboratory studies (i.e. indirect effects, predator-prey interactions). In certain cases it may still be needed to fall back on the AF-method.

3 Derivation of WFD-standards for secondary poisoning and human health

3.1 Organic compounds and metals

The text in this Chapter is focused mainly on organic compounds. Guidance focusing on metals will be presented in a separate ERL report, which is under preparation.

3.2 General approach

Before discussing the evaluation and selection of the required input data in more detail, the principles of the methodology are outlined. The assessment (if triggered, see 1.3) starts with the derivation of the concentrations in food at which no negative effects are expected for predators and humans. These concentrations are called biota-standards and denoted as $QS_{biota,\ secpois,\ fw}$, $QS_{biota,\ secpois,\ sw}$ and $QS_{biota,\ hh\ food}$.

Assuming that the trophic levels (TL) for algae, zooplankton, small fish and large fish are 1, 2, 3, and 4, respectively, the $QS_{biota,\ hh\ food}$ is considered to represent TL4 fish. This also holds for secondary poisoning of predatory birds and mammals such as seals, dolphins and seabirds at TL5 that feed on freshwater prey. For the marine environment, however, the assessment of secondary poisoning relates to a higher TL than for freshwater, since the protection goal are top predators such as killer whales and polar bears that feed on TL5-food (Figure 4). Concentrations in TL4-fish depend on the accumulation of substances from the aqueous phase by lower aquatic organisms (bioconcentration) and accumulation in the food chain from lower trophic levels to TL4 (biomagnification). These processes are represented by a laboratory bioconcentration factor (BCF) and biomagnification factors (BMF), or preferably the field bioaccumulation factors (BAF) at the appropriate trophic level.

The BCF is the ratio of the concentration in the organism (in wet weight, preferably normalised to 5% lipids [29]), divided by the water concentration. BCF values are mostly determined in the laboratory with the water phase as the exposure only exposure route. The BMF is the ratio of the concentration in the predator organism divided by the concentration in the prey organism (for hydrophobic organic chemicals commonly normalised to lipid content of prey and predator). BMF_1 describes the overall biomagnification up to larger fish (TL4) in the aquatic environment that in turn are eaten by predators (including humans). The overall BMF up to the fourth trophic level in the aquatic environment thus actually comprises three biomagnification steps. For biomagnifying substances, only the first trophic level of primary consumers is in equilibrium with the water phase. The next trophic levels deviate from equilibrium if biomagnification occurs. If biomagnification is expressed as the trophic magnification factor (TMF, which is the average increase in concentrations per trophic level) then the overall biomagnification step to TL4 is equal to TMF^3 [30,31]. For the marine environment, a second BMF_2 is included to account for

accumulation in bird and mammals at TL5 (e.g. seals, seabirds) that serve as food for top predators such as polar bears and killer whales.

The combination of BCF and BMF is represented by the bioaccumulation factor (BAF). The BAF is the ratio of the concentration in the organisms in wet weight, preferably normalised to 5% lipids [29] divided by the concentration in its surroundings (the water column). Because BAF values represent a direct measurement of bioaccumulation in the field, these BAF values at TL4 are preferred. In general, biomagnification, and thus total bioaccumulation, increases with increasing bioconcentration potential.

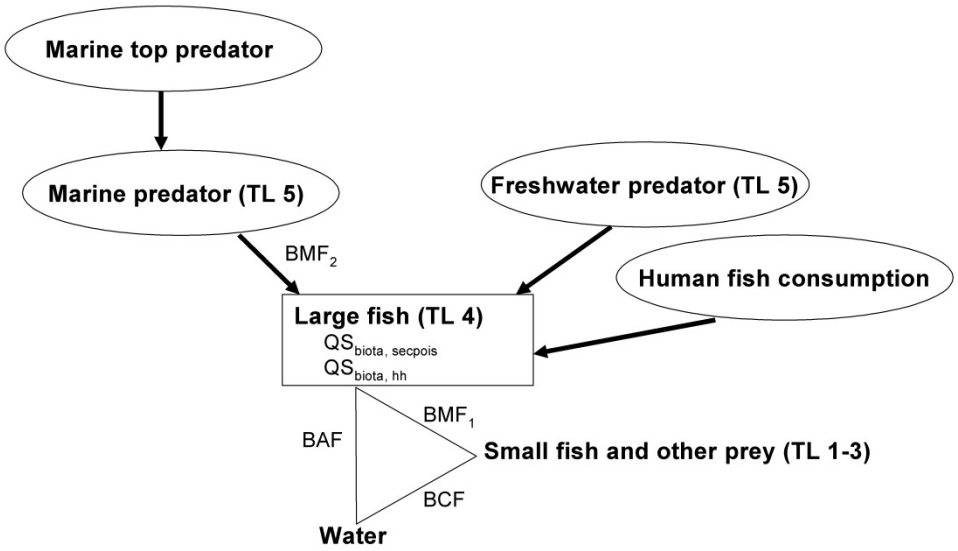

Figure 4 Scheme on how to recalculate biota standards into water concentrations. Ovals are protection goals (species to be protected); the rectangle is the trophic level on which the QS are set to protect the upper trophic levels. TL = trophic level; assuming trophic level 1 = algae; 2 = zooplankton; 3 = small fish; 4 = large fish; 5 = predatory birds, mammals, and large predatory fish. $QS_{biota,\ secpois}$ is the quality standard protecting predators (through secondary poisoning). $QS_{biota,\ hh}$ is the quality standard protecting humans (through the consumption of fish and fishery products). For $QS_{biota\ secpois,\ fw}$, (in freshwater bodies), only the BMF_1 is relevant. For $QS_{biota,\ secpois\ sw}$ in (marine waters), the BMF_1 and BMF_2 are relevant because an additional trophic level should be included to protect marine top predators. Figure copied from [32].

3.3 Collection and evaluation of bioaccumulation data

➔ Location in WFD guidance: Appendix A1.4, p. 152-156.

In principle, the evaluation of bioaccumulation data follows the general guidance on evaluation for ecotoxicity to a large extent (see ERL Report 02). All retrieved literature is read and evaluated with respect to its usefulness and reliability. The most relevant BAF and BCF studies are those performed with fish. BAF and BCF studies performed with molluscs are important for secondary poisoning as well. BAF and BCF data for other species should be carefully checked because they are prone to experimental errors. The accumulation may not reflect uptake

but adsorption to the outside of the organism. For this reason, BCF values for algae should be regarded as unreliable. A reliable BCF study should be similar in experimental set-up to the updated OECD guideline 305 [33]. At least the concentration of the (parent) compound in the aqueous phase, and in the organisms, has to be measured at several time points. No guidance exists to derive BAFs or BMFs. These data will be mostly derived from field studies. As indicated in section 3.2, field-based BAFs are preferred over the use of separate BCF- and BMF-values. For a valid BAF-study, insight into the corresponding concentrations in water at the time of organism sampling is needed. For a reliable BMF value it is necessary to know that the prey and predator species originate from the same area and from the same period in time. After evaluating a study, the results of the study are summarised by entering it into the appropriate data table (see 3.2.1).

A valid BCF/BAF ≥ 100 L/kg and/or BMF greater than 1 is used as an indication of the potential for bioaccumulation (see section 1.3). Bioaccumulation data for metals should be treated with special care, since for some metals organisms are able to regulate internal concentrations. In this case, bioconcentration and bioaccumulation may depend on the external water concentrations.

3.3.1 Data tables for bioaccumulation studies

The following sections discuss the parameters that are to be reported in the data tables, an example of which is given below. The aim is to fill the table as completely as possible. The parameters are treated in the same order as they appear in the default table, but only those parameters that are specific for bioconcentration/bioaccumulation are discussed, for general parameters, see section 2.1.1. Note that bioaccumulation studies may yield different type of results, depending on how bioconcentration is measured and expressed: BCF, BAF or BMF. Report each of these type of endpoints in separate tables and adapt the tables where necessary.

Table 10 Example of a bioconcentration data-table for freshwater organisms.

Legend to column headings	
Species properties	relevant characteristics of the test species, such as age, size, origin
Analysis method	GC = gas chromatography; MS = mass spectrometry; LSC = liquid scintillation counting; TLC = thin layer chromatography; HPLC = high performance liquid chromatography
Test type	S = static; R = renewal; F = flow through; c = closed
Purity	refers to purity of active substance or content of active substance in formulation; ag = analytical grade; tg = technical grade
Test water	am = artificial medium; dtw = dechlorinated tap water; dw = deionised/dechlorinated/distilled water; nw = natural water; rw = reconstituted water; rtw = reconstituted tap water; tw = tap water
T	Temperature
Exp. / Dep. time	exposure and depuration time
BCF-type	ethe basis of the BCF, e.g. wet weight, whole fish, edible parts
Method	method for calculation of the BCF, e.g. steady state concentrations or kinetic approach
Ri	Reliability index according to [10]. Valid studies (Ri 2 or higher) are considered for EQS-derivation, depending on relevance and considering notes on data treatment (section 1.3.4)

Species	Species properties	Analysis Method	Test type	Test compound	Purity [%]	Test water	Hardness $CaCO_3$ [mg/L]	pH	T [°C]	Exp. time	Dep. time	Exp. conc. [µg/L]	BCF [L/kg]	BCF-type	Method	Ri	Note	Ref.
Mollusca																		
Mytilus edulis	field collected, shell length 4 cm	GC-MS	F	active	99.5	am	30		10	96 h		0.5	2300	ww; edible	Corg/Cw	2	1	[a]
Pisces																		
Cyprinus carpio	8 cm	LSC	F	active, ^{14}C	99	rw		6.0-8.5	25	56 d		0.5	11000	whole fish	Corg/Cw	3	2	[b]
Oncorhynchus mykiss	1.2 g, 3% lipid	LSC, TLC, HPLC	F	active, ^{14}C	99	rw	50	7.5	21	63 d		0.005	9600	whole fish	k1/k2	2	3	[c]

Notes
1 steady state reached
2 significant mortality occurred; only information on total radioactivity
3 result based on total RA, parent confirmed by TLC and HPLC
4
5

3.3.1.1 Trophic level and lipid and/or dry weight fraction

Next to the regular species properties such as age, weight and length, two other parameter are very important for a good assessment of the bioaccumulation potential. These properties are trophic level (for field studies) and lipid fraction (in the case of hydrophobic organic chemicals) or dry weight content (for other chemicals such as metals). These parameters play an important role in normalization of the data. If trophic level is not given in the study itself, it might be derived from presented stable isotope analyses of the biota samples.

3.3.1.2 Analysis method

Similar to the toxicity data tables, a column in the bioaccumulation data tables is included that gives information on the analysis of both the aqueous phase and biological material. However, as the determination of the water and biota (or soil and biota) concentration is a prerequisite of any good bioaccumulation study, this column should give information on how the concentration is determined, not on whether the concentration is determined. Examples of such analyses are GC-FID or GC-MS (gas chromatography coupled to a flame ionisation detector or a mass spectrometer), and HPLC-UV (high-performance liquid chromatography). Especially in the case that a radiotracer is used in a BCF study, the analysis used is important. If LSC (liquid scintillation counting) is used, this means that the total radioactivity, including the parent compound and metabolites, is analysed. HPLC used in combination with radiodetection is aimed at analysis of only the parent compound.

3.3.1.3 Exposure time and depuration time

In these columns, the times of the uptake phase and, if carried out, the depuration phase are listed. As these columns refer to laboratory conditions, they are not applicable to field BAFs and BMFs.

3.3.1.4 Time of sampling

For field derived parameters it is important that samples are taken in the same time period. Therefore, the sampling time for both water and biota samples should be recorded.

3.3.1.5 Sampling area

For field studies yielding BAFs and BMFs, the location of the samples should be documented as well. For reliable BAFs and BMF, samples (water/biota, prey/predator) should be retrieved from the same area.

3.3.1.6 Exposure concentration

The concentration at which the BCF study is performed is given in this column of the BCF table. This value is important because guidelines require that the concentration meets some conditions. For example, according to the OECD guideline 305 (OECD, 1996), the highest aqueous concentration should be about one hundredth of the acute LC50 or the acute LC50 divided by an appropriate acute-to-chronic ratio, while the lowest concentration should preferably be a factor of ten below the highest concentration, but at least ten times above the limit of detection in the aqueous phase.

For BAF values, the exposure concentration is equally important. The exposure concentrations forms the basis for the BAF calculation. By tabulating the exposure concentrations any erroneously low or high value as well as a concentrations dependency of the BAF can be easily detected.

3.3.1.7 BCF, BAF or BMF

Unit: L/kg_{ww} (BCF and BAF), kg_{ww}/kg_{ww} (BMF) or kg_{lw}/kg_{lw} in case lipid normalised organism concentrations are available.

Here, the value of the BCF or BAF is denoted. The BCF value is calculated from the concentration in the organism and the water concentration. BAF values are based on the same data as the BCF, but based on field measured data. A BMF represents the ratio of concentrations of predator and prey organisms. The expression of the BCF/BAF and BMF depend on what type of measurements have been performed, e.g. normalisation to wet weight, dry weight or lipid weight. This should be explicitly indicated with a note describing the origin of the value.

BCF or BAF values used for triggering and calculating the routes of secondary poisoning and human consumption of fishery products should be whole body BCFs, expressed in L/kg. Values based on wet weight organisms are most relevant, as the organism is consumed in its "wet" form. Values expressed in dry weight organism should be recalculated to wet weight values using the data reported from the study.

It is realised that this allows for variation since these BCFs are not normalised to lipid or fat content, which dominates accumulation. ERL derivation is purely dependent on the available studies. In some (older) BCF studies, fat content may not be reported. Because there is no possibility to request studies for the purpose of ERL derivation, requirements with respect to normalisation are not applied. Including non-normalised data is preferred above excluding the data, which would possibly result in bioaccumulative substances not being triggered.

3.3.1.8 BCF/BAF/BMF type

In this column in the table, it is reported what part of the organism the BCF has been determined for. Possibilities are (e.g.): whole fish ww, whole fish dw, edible parts, non-edible parts, viscera, blood, serum, etc. For BMF, sometimes two different body parts of the prey and predator are monitored. Especially in the case when data have not been normalized special attention should be paid to the reliability of the BMF, but also for BCF and BAF, because different body parts may have different accumulation characteristics.

3.3.1.9 Method

The method that is used to calculate the BCF value is reported in this column. Basically, the method to calculate the BCF can be based on equilibrium concentrations (denoted as C_{org}/C_w) or on kinetics including the uptake and depuration rate constants (k_1 and k_2). When the BCF is determined as the quotient of the concentrations in organisms, mostly earthworms, and matrix (pore water, soil, lipid) at equilibrium, this is noted (as equilibrium). When kinetic rate constants (k_1/k_2) are

determined, the BCF is calculated as the quotient of uptake rate constant (k_1) and depuration rate constant (k_2), mostly determined independently during an uptake and a depuration phase (k_1, k_2 independent). However, in some studies, k_2 is first determined from the depuration phase and k_1 estimated from the data of the uptake phase, with this value of k_2 implied to take the non-linearity of the uptake into account (k_1 implied by fitted k_2). A further possibility is that k_1 and k_2 are fitted simultaneously by a non-linear regression model (k_1 and k_2 fitted simultaneously). The latter method might be preferred, as it takes all data together into account. In all cases, fitting could be performed on untransformed and ln-transformed data. If this is of influence on the final BCF value this should be mentioned. If the method cannot be shortly described, a reference to a note below the table can be entered here. The method is then described in more detail in the note.

3.3.1.10 Notes
Additional notes are recorded here by a number. Notes are listed below the table. The notes may include information on the analysis, a deviating basis of the BCF value (dry weight or lipid weight) or the method used to determine the BCF.

3.4 Data selection

3.4.1 Bioconcentration factor BCF

From the valid BCF studies summarised in the data tables (section 3.3), the geometric mean values per species is calculated. If (geometric mean) BCF values are available for multiple species, the geometric mean per taxon is calculated from the selected values per species. The values for fish and mussels are used for comparison with the trigger values and listed in the summary table (see section 1.7, Table 2).

When a BCF cannot be derived on the basis of experimental data, the log K_{ow} value of the compound of interest should be checked (see ERL Report 02, section 4.2.5). BCF values are needed in further ERL derivation when log $K_{ow} \geq 3$ (WFD guidance p. 17) or when derivation of an ERL for human fish consumption is triggered. When log $K_{ow} \geq 3$, calculate a BCF according to the methods in WFD guidance (section A.1.4.4.2, p. 155), which are cited from the REACH guidance [29]. These methods are briefly described below:

For substances with a log K_{ow} of 1 – 6, the following linear relationship, as developed by Veith *et al*. (1979), can be used:

$$\log BCF_{fish} = 0.85 \cdot \log K_{ow} - 0.70 \qquad (1)$$

For substances with a log K_{ow} higher than 6, a parabolic equation can be used:

$$\log BCF_{fish} = -0.20 \cdot \log K_{ow}^2 + 2.74 \cdot \log K_{ow} - 4.72 \qquad (2)$$

It should be noted that due to experimental difficulties in determining BCF values for such substances this mathematical relationship has a higher degree of uncertainty than the linear one. Both relationships

apply to compounds with a molecular weight of less than 700. For a discussion on both relationships see REACH R.7c, page 19-21 [34].

3.4.2 Biomagnification factor BMF

Experimental BMF values generally originate from field studies. Laboratory derived BMF values according to the OECD 305 test guideline cannot be used for this purpose, because these were derived in the absence of simultaneous aqueous exposure. Due to the fact that field studies are non-standard by nature, calculating a geometric mean BMF might not be justified and a value might be selected based on expert judgement. Additional information from BAF-studies may be used to select a BMF that together with the BCF would cover the BAF-values encountered in the field (see e.g. motivation in [35,36]. This final BMF is, complimentarily to the BCF and BAF, used for comparison with the trigger values and listed in the summary table (see section 1.7, Table 2).

When a BMF cannot be derived on the basis of experimental data, check the log K_{ow} value of the compound of interest, since BMF values are only needed in further ERL derivation when log $K_{ow} \geq 3$. If log $K_{ow} \geq 3$ and experimental data on BMF are not available, default BMF values will be selected, depending on the log K_{ow} of the compound of interest. The WFD guidance (citing REACH R.7.10.4.5) gives the default values for the biomagnification factors reported in Table 11 below. In this table, BMF_1 is a value for the biomagnification in the prey of predators for the freshwater environment. For the marine environment, an additional biomagnification step is included, which is reflected in the BMF_2 value. This BMF_2 is a value for biomagnification in the prey of top predators.

The most relevant values for BMF_1 are those for biomagnification from small to larger fish (either fresh or marine water). These larger fish then serve as food for predators such as otters and herons, and seals in the marine environment. However, as fish at trophic level 4 is 3 three levels above the trophic level that is in equilibrium with the water phase, BMF_1 should thus also include three trophic magnification steps. Such a BMF does not represent a single predator-prey relationship. Besides that, such a BMF including three trophic levels will not be available for a BMF from fish to fish. Data for biomagnification from other small species such as crustaceans to fish might be useful as well, but care must be taken that in the further assessment of secondary poisoning, BCF and BMF values are in accordance with each other. If no reliable estimate of the BAF at trophic level 4 can be generated, an alternative might be to use the trophic magnification factor instead of a BMF. To account for magnification over three trophic levels, the value TMF^3 could be used. Another group of prey that might be relevant to the route of secondary poisoning are mussels. If mussels are directly consumed by birds or mammals and a BAF value for mussels is available, a biomagnification step would be absent. However, there are also several common fish species that feed on mussels. In such a case BMF data on accumulation from mussels to fish would be relevant.

For the marine environment another biomagnification step is considered by introducing the BMF_2 value. This step refers to the biomagnification from fish to small mammals and birds. For the marine environment, a

good example is the biomagnification from fish to seals. The latter species then serve as prey for top predators such as polar bears and killer whales. However, besides data for the marine environment, other data for biomagnification from fish to fish-eating birds and mammals should be considered as well.

If no reliable data for biomagnification are available, the default values from Table 11 can be used, which originates from [37]. Column 1 of Table 11 shows (ranges of) log K_{ow} values. If one or more experimental BCF data are available, the K_{ow} values from the tables are not needed. The BMF is determined according to the BCF ranges in the second column. If there is no experimental BCF value, the BCF is estimated from log K_{ow} (see 3.4.1), and BMF is derived according to the log K_{ow} ranges in the first column of Table 11.

Table 11 Default BMF-values for organic substances.

log Kow of substance	**BCF (fish)**	**BMF1**	**BMF2**
< 4.5	< 2000	1	1
4.5 - < 5	2000-5000	2	2
5 – 8	> 5000	10	10
> 8 – 9	2000 – 5000	3	3
> 9	< 2000	1	1

3.4.3 Bioaccumulation Factor BAF

As is apparent from the above, the derivation of standards for secondary poisoning and human health should be based on a comprehensive evaluation of BCF, BMF and BAF-values. In general, preference is given to the use of BAFs instead of using the product of BCF and BMF, because the BAF is based on field samples and includes all possible uptake routes [32]. For a valid BAF, however, insight into the corresponding concentrations in water is needed and the BAF should be valid for the appropriate TL. This can for example be done by a regression of BAF values as a function as trophic level. Examples of such regression can be found in the derivation of a BAF at TL4 for hexachlorobenzene [32] or mercury [38]. Depending on the type and validity of information, it sometimes may be more appropriate to rely on the combination of BCF and BMF.

3.4.4 BCF and BAFs for metals

Many organisms can keep their body concentration of metal relatively constant within certain concentration range, while the water concentration varies. Variation in BCF or BAF is then not caused by accumulation but by regulation. Inverse relationships of BCF/BAF with external water concentration have been observed [5,39]. The BCF concept as applicable to many organic substances is not valid for metals and BCFs (BAFs) for metals (BAF) values for metals should not be used in the same way, nor can they be simply averaged. If a relevant relationship between BAF and external water concentration is observed BAf values derived in this way should be preferred. An example of this is the use of BAF values for the derivation of a QS for uranium in water for the protection goal secondary poisoning [40].

➔ WFD guidance deals with the use of BCF values for metals in:
section 2.4.3.1, page 18 and 19.
section 4.7.2.2, on pages 88 to 91.

The following aspects should be taken into account,

- Is the metal essential or non-essential?
- Concentrations in BCF studies should be well below toxicity levels.
- BAF studies are preferred over BCF studies.
- Investigate the relationship between internal and external concentration for various organisms and conclude whether averaging of BCFs or BAFs is allowed.
- Where averaging is not allowed, follow the scheme on page 90 of the WFD guidance for use of the BCF (BAF) – external concentration relationship in QS derivation.

3.5 Derivation of biota standards, calculation to water

If data allow, the methodology as presented in ERL report 7 should be used. If this is not possible, the default method according to the WFD-guidance is applied, which is summarised below.

3.5.1 Secondary poisoning

The biota-based standards for secondary poisoning are calculated from the lowest NOEC for birds or mammals (see ERL Report 02, Chapter 6) divided by the appropriate assessment factor.

Table 12. Assessment factors for extrapolation of mammalian and bird toxicity data.

TOXoral	**Duration of test**	**AForal**
$LC50_{bird}$	5 days	3000
$NOEC_{bird}$	Chronic	30
$NOEC_{mammal,\ food\_chr}$	28 days 90 days chronic	300 90 30

➔ Location in WFD guidance: Section 4.4.4.1, p. 71.

In general, the lowest value should be selected, but if for a single species more data are available from studies with different durations, preference is given to studies with a longer test duration. However, if the NOEC from a short-term study is lower than the NOEC of a longer-term study, the lowest endpoint should be used, but it may then be combined with the lower assessment factor that is normally associated with the longer-term study (e.g. combine a 28-days NOEC with an AF of 30). Although gestation studies relate to short-term exposure, it is advised to use an assessment factor of 90, because these studies are performed during a very sensitive part of the life-cycle.

If data for multiple species are available, statistical extrapolation may be considered. No explicit guidance is available on the number of species required, but analogues to direct ecotoxicity this would imply a number of at least 10 species.

In the WFD-guidance, the biota standards are calculated for the TL to which they apply, i.e. TL4 for the freshwater compartment and TL5 for the saltwater compartment. In other words, the correction for trophic level is made already at the level of the biota standard. Following this reasoning, in case biomagnification is relevant (BMF_2 > 1), the derived biota standards for freshwater and saltwater differ (see Equation 3 and 4)

$$QS_{biota,secpois,fw} = \frac{TOX_{oral}}{AF} \quad (3)$$

$$QS_{biota,secpois,sw} = \frac{TOX_{oral}}{AF \times BMF_2} \quad (4)$$

The $QS_{fw,\ secpois}$ and $QS_{sw,\ secpois}$ are then calculated from the biota standard by dividing the respective QS_{biota}-values by the product of BCF and BMF_1:

$$QS_{fw,secpois} = \frac{QS_{biota,\ secpois,\ fw}}{BCF \times BMF_1} \quad (5)$$

$$QS_{sw,\ secpois} = \frac{QS_{biota,\ secpois,\ sw}}{BCF \times BMF_1} \quad (6)$$

Confusingly, the Priority Substances Directive 2013/39/EU lists only one biota standard, which is valid for freshwater and saltwater. This means that the correction for trophic level should be made upon compliance check by introducing a correction factor when recalculating biota standards into corresponding concentrations in water. The $QS_{fw,\ secpois}$ and $QS_{sw,\ secpois}$ are then calculated from a single biota standard by dividing the $QS_{biota,\ secpois}$ at TL4 by the product of BCF and BMF_1 for freshwater (Equation 7) and by the product of BCF, BMF_1 and BMF_2 for marine waters (Equation 8).

$$QS_{fw,secpois} = \frac{QS_{biota,\ secpois}}{BCF \times BMF_1} \quad (7)$$

$$QS_{sw,\ secpois} = \frac{QS_{biota,\ secpois}}{BCF \times BMF_1 \times BMF_2} \quad (8)$$

The term BCF x BMF may be replaced by a bioaccumulation factor (BAF), which is the ratio of the concentration in the organisms (in wet weight, preferably normalised to 5% lipids) divided by the concentration in its surroundings (the water column). The BAF is determined from field samples and includes both uptake from the water phase and uptake via food. In this case, care should be taken that the BAF is derived for the appropriate trophic level, i.e. TL4 for freshwater (Equation 9), and TL5 for saltwater (Equation 10). If for saltwater a BAF at TL4 is used, this should be corrected to TL5 using the BMF.

$$QS_{fw,\ secpois} = \frac{QS_{biota,\ secpois}}{BAF_{TL4}} \quad (9)$$

$$QS_{sw,\ secpois} = \frac{QS_{biota,\ secpois}}{BAF_{TL5}} \quad (10)$$

BMFs are generally derived from field studies, which nowadays often study the transfer of a compound through the food chain as a function of trophic level. In that case, the BMF is referred to as Trophic Magnification Factor (TMF).

3.5.2 Human health

The derivation of the $QS_{biota,\ hh\ food}$ is performed according to the WFD-guidance.

➔ Location in WFD guidance: Section 4.5, p. 82.

The $QS_{biota,\ hh\ food}$ is the concentration in fish which leads to an intake of a compound of at most 10% of the human toxicological risk limit (TTL_{hh}), given a body weight of 70 kg, and a daily fish consumption of 115 g per day. For derivation of the TTL_{hh}, see (see ERL rapport 02, Chapter 7). Because it is assumed that freshwater and saltwater fish eaten by humans are from the same TTL, the calculation is the same for freshwater and saltwater.

$$QS_{biota,hh\ food} = \frac{0.1 \times 70 \times TTL_{hh}}{0.115} \quad (11)$$

This biota standard is converted into a corresponding water concentration using the product of BCF and BMF_1, or an experimental BAF at TL4 (TL = trophic level).

$$QS_{water,\ hh food} = \frac{QS_{biota,\ hh\ food}}{BCF \times BMF_1} \quad (12)$$

$$QS_{water,\ hh\ food} = \frac{QS_{biota,\ hh\ food}}{BAF_{TL4}} \quad (13)$$

3.6 Quality standards for freshwater for drinking water abstraction

The derivation of this route follows the methodology described in the WFD-guidance. If no published EU or WHO-drinking water standard is available, a provisional standards fro the TTL_{hh} according to the formulas in the WFD-guidance. It should also be checked if a drinking water limit has been derived for specific purposes within other national frameworks.

➔ Location in WFD guidance: Section 3.9, p. 71.

4 Derivation of national risk limits SRC and NC

4.1 Serious Risk Concentration for direct ecotoxicity

See ERL Report 01, section 4.6 for general guidance on the SRC_{eco}. The SRC_{eco} is the geometric mean of all available chronic toxicity data (that have been judged valid and have been compiled in the aggregated data table). If not enough chronic toxicity data are available, the SRC_{eco} is calculated as the geometric mean of all (aggregated) acute data, divided by an assessment factor of 10. The two values are compared and the lowest value is selected as SRC_{eco}.

The aggregated data tables with acute and chronic aquatic toxicity data are used for the derivation of the SRC_{eco} according to the assessment factor scheme in Table 13. In case a pooled data set for freshwater and marine toxicity data is used for QS derivation (see section 2.3), the pooled (aggregated) data set is also used for SRC derivation. In this case, one $SRC_{water, eco}$ is derived that is valid for both the freshwater and the marine compartment. No additional assessment factor is used for derivation of the $SRC_{sw, eco}$. When the freshwater and marine data have not been pooled for QS derivation, the assessment factor scheme in Table 13 is applied to the separate freshwater and marine aggregated data sets to derive an $SRC_{fw, eco}$ and $SRC_{sw, eco}$.

In addition, take account of the following:

- In principle, an acute-to-chronic ratio (ACR) of 10 is applied to the acute toxicity data to compare acute L(E)C50s with chronic NOECs (or EC10s). One may deviate from this factor of 10 if more information on the ACR for the specific compound or endpoint is available [41].
- For the aquatic compartment, comparison between chronic data and acute data is not performed when chronic data are available for *at least* three species, which should represent the three specified trophic levels from the base set of REACH guidance: algae, *Daphnia* and fish (see Table 13).
- When the SRC_{eco} is to be reported with confidence limits, the computer program *ETX* 2.1 [42] is used to calculate the median HC_{50} and its 90% confidence interval. The HC_{50} is equal to the geometric mean of log-normally distributed toxicity data.
- The SRC_{eco} is always taken as the geometric mean of (either acute or chronic) toxicity data, irrespective of whether these data are log-normally distributed or not. If the data from which the SRC_{eco} is calculated do not fit a normal distribution, it suffices to note this briefly in the report section where the SRC_{eco} derivation is presented.
- For metals the added risk approach should be followed. The SRC_{eco} is defined as the background concentration plus the serious risk addition (SRA_{eco}).

Table 13. Assessment factors used to derive the SRC_{eco} for the aquatic compartment.

Available test results	**Additional criteria**	**SRC_{eco} based on**	**Assessment factor**
only L(E)C50s and no NOECs		geometric mean of L(E)C50s	10
1 NOEC[1]	**none** of three specified taxa[2] is represented	geometric mean of L(E)C50s	10
1 NOEC[1]	one of three specified taxa[2] is represented AND geometric mean of L(E)C50s / 10 < NOEC value	geometric mean of L(E)C50s	10
1 NOEC[1]	one of three specified taxa[2] is represented AND geometric mean of L(E)C50s / 10 ≥ NOEC value	NOEC value	1
≥ 2 NOECs[1]	**none** of three specified taxa[2] is represented	geometric mean of L(E)C50s	10
≥ 2 NOECs[1]	one or two of three specified taxa[2] is represented AND geometric mean of L(E)C50s / 10 < geometric mean[3] of NOECs	geometric mean of L(E)C50s	10
≥ 2 NOECs[1]	one or two of three specified taxa[2] is represented AND geometric mean of L(E)C50s / 10 ≥ geometric mean[3] of NOECs	geometric mean[3] of NOECs	1
≥ 3 NOECs[1]	≥ 3 of three specified taxa[2] are represented	geometric mean[3] of NOECs	1

1: this may also be an EC10 value.
2: the 3 taxa for which NOEC data (and/or EC10 values) should be available are **algae, *Daphnia* and fish**.
3: the geometric mean of all available NOECs (and/or EC10 values) is calculated; including the values that do not belong to the specified taxa.

4.2 Negligible Concentration

The Negligible Concentration for fresh- and saltwater (NC_{fw} and NC_{sw}) are calculated by dividing the respective AA-EQS_{fw} and AA-EQS_{sw} by a factor of 100, except when in the case of genotoxic carcinogens the final AA-EQS is based on human fish consumption (see section 1.5). In that case, the NC is not derived.

5 References

1. EC. 2011. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Guidance Document No. 27. Technical Guidance For Deriving Environmental Quality Standards. Brussels, Belgium. European Commission. Report Technical Report - 2011 - 055.
2. Brock TCM, Arts GHP, Ten Hulscher TEM, De Jong FMW, Luttik R, Roex EWM, Smit CE, Van Vliet PJM. 2011. Aquatic effect assessment for plant protection products; Dutch proposal that addresses the requirements of the Plant Protection Product Regulation and Water Framework Directive. Wageningen, the Netherlands. Alterra. Report 2235.
3. Smit CE, Arts GHP, Brock TCM, Ten Hulscher TEM, Luttik R, Van Vliet PJM. 2013. Aquatic effect and risk assessment for plant protection products; Evaluation of the Dutch 2011 proposal. Wageningen, the Netherlands. Alterra Wageningen UR (University & Research centre). Report 2463.
4. Brix KV, DeForest DK. 2000. Critical review of the use of bioconcentration factors for hazard classification of metals and metal compounds. Kirkland, WA, USA. Report 555-3690-001.
5. McGeer JC, Brix KV, Skeaff JM, DeForest DK, Brigham SI, Adams WJ, Green A. 2003. Inverse relationship between bioconcentration factor and exposure concentration for metals: implications for hazard assessment of metals in the aquatic environment. Environmental Toxicology and Chemistry 22(5), 1017-1037.
6. Ikemoto T, Tu NPC, Okuda N, Iwata A, Omori K, Tanabe S, Tuyen BC, Takeuchi I. 2008. Biomagnification of trace elements in the aquatic food web in the Mekong delta, South Vietnam using stable carbon and nitrogen isotope analysis. Archives of Environmental Contamination and Toxicology 54, 504-515.
7. EC. 2008. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Communities L 353, 1-1355.
8. EC. 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Official Journal of the European Communities L 327, 1-72.
9. Traas TP. 2001. Guidance document on deriving environmental risk limits. Bilthoven, The Netherlands. National Institute for Public Health and the Environment. Report 601501012.
10. Klimisch HJ, Andreae M, Tillman U. 1997. A systematic approach for evaluating the quality of experimetnal toxicological and ecotoxicological data. Regulatory Toxicology and Pharmacology 25, 1-5.
11. ECHA. 2008. Guidance on information requirements and chemical safety assessment. Chapter R.10: Characterisation of dose

[concentration]-response for environment. Helsinki, Finland. European Chemicals Agency.
12. Vethaak AD, Jol JG, Martínez-Gómez C. 2011. Effects of cumulative stress on fish health near freshwater outlet sluices into the sea: A case study (1988-2005) with evidence for a contributing role of chemical contaminants. Integrated Environmental Assessment and Management 7(3), 445-458.
13. Cairns Jr J. 1986. The myth of the most sensitive species. Bioscience 36(10), 670-672.
14. Dang Z, Smit CE. 2008. Environmental risk limits for carbendazim. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 601716014.
15. Moermond CTA, Van Vlaardingen PLA, Vos JH, Verbruggen EMJ. 2008. Environmental risk limits for dimethoate. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 601714001.
16. Moermond CTA, Vos JH, Verbruggen EMJ. 2008. Environmental risk limits for organophosphorous pesticides. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 601714004 (with erratum).
17. Smit CE. 2014. Water quality standards for imidacloprid. Proposal for an update according to the Water Framework Directive. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 270006001.
18. Van Leeuwen LC, Vos JH, Mensink BJWG. 2008. Environmental risk limits for lambda-cyhalothrin. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 601716001.
19. Brock TCM, Alix A, Brown CD, Capri E, Gottesbüren BFF, Heimbach F, Lythgo CM, Schulz R, Streloke M, Linking aquatic exposure and effects: risk assessment of pesticides. 2010, Boca Raton, USA: SETAC Press & CRC Press, Taylor & Francis Group.
20. Campbell PJ, Arnold D.J.S., Brock TCM, Grandy NJ, Heger W, Heimbach F, Maund SJ, Streloke M. 1999. Guidance document on higher-tier aquatic risk assessment for pesticides (HARAP). Brussels, Belgium. SETAC-Europe.
21. Giddings JM, Brock TCM, Heger W, Heimbach F, Maund SJ, Norman SM, Ratte HT, Schäfers C, Streloke MS. 2002. Community-level Aquatic System Studies: Interpretation Criteria (CLASSIC) Proceedings from the CLASSIC workshop. Schmallenberg, Germany / Pensacola, USA. Fraunhofer Institute/SETAC.
22. OECD. 2006. Guidance document on Simulated Freshwater Lentic Field tests (outdor microcosms and mesocosms). Series on Testing and Assessment, No 53. OECD Environment Directorate, Paris.
23. De Jong FMW, Brock TCM, Foekema EM, Leeuwangh P. 2008. Guidance for summarising and evaluating aquatic micro- and mesocosm studies. A guidance document of the Dutch Platform for the Assessment of Higher Tier Studies. Bilthoven, The Netherlands. National Institute for Public Health and the Environment. Report 601506009.

24. EFSA. 2013. Scientific Opinion. Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Journal 11(7), 3290.
25. De Jong FMW, Bakker FM, Brown K, Jilesen CJTJ, Posthuma-Doodeman CJAM, Smit CE, Van der Steen JJM, Van Eekelen GMA. 2010. Guidance for summarising and evaluating field studies with non-target arthropods: A guidance document of the Dutch Platform for the Assessment of Higher Tier Studies. Bilthoven, Nederland. National Institute for Public Health and the Environment. Report 601712006.
26. De Jong FMW, Van Beelen P, Smit CE, Montforts MHMM. 2006. Guidance for summarising earthworm field studies. A guidance document of the Dutch Platform for the Assessment of Higher Tier Studies. Bilthoven, The Netherlands. National Institute for Public Health and the Environment. Report 601506006.
27. Brock TCM, Lahr J, Van den Brink PJ. 2000. Ecological risks of pesticides in freshwater ecosystems. Part 1. Herbicides. Wageningen, the Netherlands. Alterra Green World Research. Report 088.
28. Brock TCM, Van Wijngaarden RPA, Van Geest GJ. 2000. Ecological risks of pesticides in freshwater ecosystems. Part 2: Insecticides. Wageningen, the Netherlands. Alterra Green World Research. Report 089.
29. ECHA. 2014. Guidance on information requirements and chemical safety assessment. Chapter R.7a: Endpoint specific guidance. Version 3.0. Helsinki, Finland. European Chemicals Agency. Report ECHA-14-G-03-EN.
30. Burkhard LP, Borga K, Powell DE, Leonards P, Muir DCG, Parkerton TF, Woodburn KB. 2013. Improving the quality and scientific understanding of trophic factors (TMFs). Environmental Science & Technology 47, 1186-1187.
31. Verbruggen EMJ. 2014. New method for the derivation of risk limits for secondary poisoning. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 2014-0097.
32. Moermond CTA, Verbruggen EMJ. 2012. An evaluation of bioaccumulation data for hexachlorobenzene to derive water quality standards according to the EU-WFD methodology. IEAM 9(1), 87-97.
33. OECD. 1996. OECD Guideline for the Testing of Chemicals: 305. Bioconcentration: Flow-through Fish Test, 23. Paris, France. Organisation for Economic Cooperation and Development (OECD).
34. ECHA. 2014. Guidance on information requirements and chemical safety assessment. Chapter R.7c: Endpoint specific guidance. Version 2.0. Helsinki, Finland. European Chemicals Agency. Report ECHA-14-G-06-EN.
35. Moermond CTA, Verbruggen EMJ, Smit CE. 2010. Environmental risk limits for PFOS. A proposal for water quality standards in accordance with the Water Framework Directive. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 601714013.
36. Smit CE, Posthuma-Doodeman CJAM, Verbruggen EMJ. 2012. Environmental risk limits for octamethylcyclotetrasiloxane in

water. A proposal for water quality standards in accordance with the Water Framework Directive. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 601714020.

37. ECHA. 2012. Guidance on information requirements and chemical safety assessment. Chapter R.16: Environmental Exposure Estimation. Version 2.1. Helsinki, Finland. European Chemicals Agency. Report ECHA-10-G-06-EN.
38. Verbruggen EMJ, Van Herwijnen R, Smit CE. 2015. Derivation of a water-based quality standard for secondary poisoning of mercury. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 2015-0058.
39. DeForest DK, Brix KV, Adams WJ. 2007. Assessing metal bioaccumulation in aquatic environments: The inverse relationship between bioaccumulation factors, trophic transfer factors and exposure concentration. Aquatic Toxicology 84, 236-246.
40. Van Herwijnen R, Verbruggen EMJ. 2014. Water quality standards for uranium : Proposal for new standards according to the Water Framework Directive. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 270006003.
41. Verbruggen EMJ, Posthumus R, Van Wezel AP. 2001. Ecotoxicological serious risk concentrations for soil, sediment and (ground)water: updated proposals for first series of compounds. Bilthoven, the Netherlands. National Institute for Public Health and the Environment. Report 711701020.
42. Van Vlaardingen P, Traas TP, Wintersen AM, Aldenberg T. 2004. ETX 2.0. A program to calculate hazardous concentrations and fraction affected, based on normally distributed toxicity data. Bilthoven, The Netherlands. National Institute for Public Health and the Environment. Report 601501028/2004.
43. ECHA. 2014. Guidance on information requirements and chemical safety assessment. Chapter R.7b: Endpoint specific guidance. Version 2.0. Helsinki, Finland. European Chemicals Agency. Report ECHA-14-G-05-EN.
44. Maltby L, Arnold D, Arts G, Davies J, Heimbach F, Pickl C, Poulsen V, Aquatic Macrophyte Risk Assessment for Pesticides. 2010, Boca Raton, USA: SETAC Press & CRC Press, Taylor & Francis Group.

List of abbreviations

AA-EQS	annual average environmental quality standard
ACR	acute to chronic ratio
ADI	acceptable daily intake
AF	assessment factor
ag	analytical grade
am	artificial medium
AMA	amphibian metamorphosis assay
BAF	bioaccumulation factor
BCF	bioconcentration factor
BMF	biomagnification factor
bw	body weight
c	closed (exposure) system
CAS	chemical abstract service
CD	commission directive
CF	continuous flow system
c.i.	confidence interval
CMR	carcinogenic, mutagenic, reprotoxic
d	days
DT50	half life time for dissipation of a substance from an environmental compartment
dtw	dechlorinated tap water
dw	de-ionised water, dechlorinated water or distilled water dry weight
DW	drinking water
DWQG	drinking-water quality guidelines
DWS	drinking-water standard
EC	European commission; effect concentration
ECx	effect concentration at which an effect of x% is observed, generally EC10 and EC50 are calculated
ECHA	European Chemicals Agency
EEC	European economic community (replaced by EU)
EFSA	European Food Safety Authority
ELS	early life stage
EqP	equilibrium partitioning
EQS	environmental quality standard
ERL	environmental risk limit
EU	European union
F	flow through system
FHI	Fraunhofer Institute
FID	flame ionisation detection
FSDT	fish sexual development test
FSTRA	fish short term reproduction assay
GC	gas chromatography
h	hours
HCx	hazardous concentration at which x percent of species is potentially affected
HPLC	high pressure liquid chromatography
IenM	Dutch Ministry of Infrastructure and Environment
IF	intermittent flow system

INS	International and National Environmental Quality Standards for Substances in the Netherlands (In Dutch: (Inter)nationale Normen Stoffen)
ISO	international organisation for standardisation
LCx	effect concentration at which x% lethality is observed, generally LC50 and LC10 are calculated
lg	laboratory grade
LSC	liquid scintillation counting
LOEC	lowest observed effect concentration
MAC	maximum acceptable concentration
MAC-EQS	maximum acceptable concentration-environmental quality standard
MDD	Minimum Detectable Difference
min	minutes
mo	months
MPC	maximum permissible concentration
MS	mass spectrometry, Microsoft™
NA	negligible addition
NC	negligible concentration
NOAEL	no observed adverse effect level
NOEC	no observed effect concentration
NOEL	no observed effect level
nw	natural water, such as lake water, river water, sea water, well water
OECD	organisation for economic co-operation and development
pa	pro analyse
PRC	principal response curve
PPP	plant protection product
ppt	parts per thousand or parts per trillion
psu	practical salinity unit
QS	quality standard
QSAR	quantitative structure activity relationship
R	renewal system
REACH	Registration, Evaluation, Authorisation and Restriction of Chemical substances
rg	reagent grade
rtw	reconstituted tap water: tap water with additional salts
rw	reconstituted water: (natural) water with additional salts
RIVM	national institute for public health and the environment
S	static
Sc	static, closed system
sp.	species
SRA_{eco}	ecotoxicological serious risk addition
SRC	serious risk concentration
SRC_{eco}	ecotoxicological serious risk concentration
susp	suspended particulate matter
SSD	species sensitivity distribution
TDI	tolerable daily intake
tg	technical grade
TGD	Technical Guidance Document
TL	trophic level in secondary poisoning assessment and biomagnification studies
TLC	thin layer chromatography
TTL_{hh}	toxicological threshold level for human health

TLm	median tolerance limit; also encountered as: median threshold limit
TMF	trophic magnification factor
tw	tap water
TWA	time weighted average
UV	ultraviolet
w	weeks
WFD	water framework directive
WHO	world health organisation
ww	wet weight
y	years

Appendix 1. Established guidelines for aquatic ecotoxicity tests

Aquatic organisms
➔ Location in WFD guidance: Appendix A1.3.2.10, p. 136-138.
This section has been updated with respect to the version in the WFD guidance.
International guidelines for performing aquatic toxicity studies exist for many species. The most frequently used guidelines for laboratory studies are summarised, grouped according to taxon

Algae and cyanobacteria
OECD 201. Alga, growth inhibition test. Applicable to several species of green algae, cyanobacteria and diatoms. The EC50 from this 72-h algae test is considered an acute value, the NOEC or EC10 a chronic value. The previous version from 1984 mentions both growth rate and biomass (sometimes called growth, calculated as area under the curve) as endpoints. The guideline was revised in 2011. In the 2011-version, the biomass integral is no longer included and growth rate is the preferred response variable. Yield (increase in cell numbers) is included to meet regulatory demands in some countries. In line with the REACH- and WFD-guidance, growth rate is the preferred endpoint for ERL-derivation (REACH R.7b [43], section 7.8.4.1). However, if only biomass is presented, this value can be used as well. The result for the endpoint biomass (area under the curve) is generally somewhat lower than the result for growth rate and can therefore be considered as a conservative value.

Crustacea
OECD 202. Daphnia sp., acute immobilisation test. For the derivation of EQSs for water, only the EC50 from this 48-h acute toxicity study is considered. The endpoint is immobility, as indicated by the inability to swim after agitation, or mortality.

OECD 211. *Daphnia magna* reproduction test. This is a chronic test with water fleas. The most important endpoint is the number of young per female (both young and parent alive). Other endpoints are the survival of the parent animals and time to production of first brood. Additionally, parameters such as growth (e.g. length) of the parent animals, and the intrinsic rate of population increase are useful endpoints.

Insecta
OECD 235. *Chironomus* sp., acute immobilisation test. Water only test with *Chironomus* sp., based on OECD 202. *C. riparius* is the preferred species, *C. dilutus* and *C. yoshimatsui* are mentioned as alternative species. Endpoint is immobility after 48 hours.

OECD 219. Sediment-water chironomid toxicity test using spiked water. This test is similar to OECD guideline 218 (see 5.4). Endpoints from this test can only be used for aquatic ERLs if it is possible to express the endpoint on the basis of concentrations in the water phase during exposure. For most substances, using the standard OECD artificial

sediment will result in decline of water concentrations in time. Therefore, modifications of the test by using inert substrate (quartz sand, sheet cloth) is preferred. These are indicated as an option in guideline 233.

OECD 233. Sediment-water chironomid life-cycle toxicity test. This test is an extension of OECD 218 and 219 (see above) and covers the early part of the 2nd generation. Measured endpoints are the total number of adults emerged (for both 1st and 2nd generations), development rate (for both 1st and 2nd generations), sex ratio of fully emerged and alive adults (for both 1st and 2nd generations), number of egg ropes per female (1st generation only) and fertility of the egg ropes (1st generation only). If effects can be expressed on the basis of concentrations in the water phase over the duration of the test, the results can be used as a basis for aquatic ERLs. Note that this can only be done when based on actual analyses of the water phase during the experiment; data from comparable systems, e.g. from a water/sediment degradation study can serve as confirmatory information, but cannot be used as a sole basis for extrapolation of initial concentrations in water to corresponding values throughout the test.

OECD test guidelines for a two-generation test with mysid shrimps (*Americamysis bahia*) and for development and reproduction of the copepod *Amphiascus tenuiremis* are currently under development.

Pisces
OECD 203. Fish, acute toxicity test. For the derivation of EQSs for water, only the LC50 from this 96-h acute toxicity study is considered. The recorded endoint is mortality.

OECD 204. Fish, prolonged toxicity test. 14-day Study. This study is also considered as an acute toxicity study, and consequently, in most cases, only the LC50 is used for the derivation of EQSs.

OECD 210. Fish, early-life stage toxicity test. This test with fish is a chronic test which covers the life cycle from eggs to free feeding juvenile fish. The recorded endpoints are mortality at all stages, time to hatch, hatching success, length, weight and any morphological or behavioural abnormalities.

OECD 212. Fish, short-term toxicity test on embryo and sac-fry stages. In the guideline it is stated that this test can be used as a screening test for chronic toxicity. Especially for species that cannot be kept under laboratory circumstances for a period long enough to perform a full early-life stage (ELS) test, this test can be a useful alternative. Because the sensitive life stages from egg to sac-fry are covered in this test, it can be considered a chronic test. However, it is expected to be less sensitive than the full ELS test. The same endpoints are recorded as for the full ELS test.

OECD 215. Fish, juvenile growth test. Because the recorded endpoint is growth during 28 days and the criterion is the NOEC or EC10, the test can be regarded as chronic.

OECD 229. Fish short term reproduction assay (FSTRA). This is an *in vivo* screening assay where sexually mature male and spawning female fish are held together and exposed to a chemical during a limited part of their life-cycle (21 days). At termination of the 21-day exposure period, vitellogenin and secondary sexual characteristics are measured in males and females as indicators of endocrine activity of the test chemical. Additionally, quantitative fecundity (egg production) is monitored daily throughout the test. In view of the duration and endpoints, this test is considered as chronic. Effects on egg production can be used for ERL-derivation, vitellogenin, secondary sexual characteristics and gonadal histopathology may be used as additional information. The short term reproduction assay was validated in the fathead minnow (*Pimephales promelas*) and this is the recommended species.

OECD 230. 21-day Fish assay: A short-term screening for oestrogenic and androgenic activity, and aromatase inhibition. This protocol describes an *in vivo* screening assay for certain endocrine active substances where sexually mature male and spawning female fish are held together and exposed to a chemical during a limited part of their life-cycle (21 days). This assay covers the screening of oestrogenic and androgenic activity, and aromatase inhibition. The assay was validated on the fathead minnow (*Pimephales promelas*), the Japanese medaka (*Oryzias latipes*) and the zebrafish (*Danio rerio*); however zebrafish does not allow the detection of androgenic activity. At termination of the 21-day exposure period, vitellogenin and/or secondary sexual characteristics are measured in males and females. This test is comparable to OECD 229, but the latter also measures actual fecundity and gonadal histopathology for the fathead minnow. Results may be considered as additional information for ERL-derivation.

OECD 234. Fish (FSDT). This test protocol is in principle an enhancement of OECD 210: Fish, Early Life Stage Toxicity Test, where the exposure is continued until the fish are sexually differentiated. The test is validated for Japanese medaka (*Oryzias latipes*), zebrafish (*Danio rerio*) and three spined stickleback (*Gasterosteus aculeatus*), and partially validated for fathead minnow (*Pimephales promelas*). The FSDT assesses early life-stage effects and potential adverse consequences of putative endocrine disrupting chemicals (e.g. oestrogens, androgens and steroidogenesis inhibitors) on sexual development. By combining two core endocrine endpoints, vitellogenin (VTG) concentration and phenotypic sex ratio, the mode of action of the test chemical can be indicated. According to the guideline, the FSDT can be used for hazard and risk assessment because the change in phenotypic sex ratio is a population-relevant parameter. For stickleback, however, this endpoint should not be used because the validation data available so far showed uncommon alterations of phenotypic sex ratio.

OECD 236. Fish (FET) Test. Test to determine the acute toxicity or lethality on fish embryonic stages. Endpoints are (i) coagulation of fertilised eggs, (ii) lack of somite formation, (iii) lack of detachment of the tail-bud from the yolk sac, and (iv) lack of heartbeat. At the end of the exposure period, acute toxicity is determined based on a positive outcome in any of the four apical observations recorded, and the LC50 is

calculated. Although this is a short-term test, the results can be used as chronic endpoints in view of the life-stage examined.
An OECD guideline for a multi-generation test with medaka (*Oryzias latipes*) is currently under development.

Macrophyta
OECD 221. *Lemna* sp. growth inhibition test. For this 7-d test with duckweed the same considerations can be made as for the algal test (OECD 201): the EC50 from this test is considered an acute value, the NOEC or EC10 a chronic value. Both chronic and acute data should be retrieved from the test. The preferred endpoints are growth rate (based on frond number) or biomass (dry weight, fresh weight or frond area). Extensive information on testing and assessment of aquatic macrophytes is provided by the AMRAP workshop [44]. Based on the recommendations of this workshop, test guidelines with the aquatic macrophyte *Myriophyllum spicatum* in water-only and water/sediment systems are currently under development within OECD.

Amphibia
OECD 231. The amphibian metamorphosis assay (AMA). The Amphibian Metamorphosis Assay (AMA) is a screening assay intended to empirically identify substances which may interfere with the normal function of the hypothalamic-pituitary-thyroid (HPT) axis. The assay was validated with the species *Xenopus laevis*, which is the recommended species. The assay has a duration of 21 days, endpoints are mortality, developmental stage, weight, snout-to-vent length and hind limb length. Histopathology of the thyroid gland is included. Mortality, weight and developmental stage are considered as chronic endpoints suitable for ERL-derivation, the other endpoints may be used as additional information.