This project is funded by
the Horizon 2020 Framework
Programme of the European Union

EuroMix

European Test and Risk Assessment Strategies for Mixtures

Project number 633172

Collaborative project
H2020-SFS-2014-2

Deliverable D8.3 – Report

Guidance on the use of the new strategy for tiered testing and assessment

WP 8 – **Scientific integration and guidance for tiered risk assessment**

Due month of deliverable: Month 45
Actual submission month: Month 48

Deliverable leader: KI

Document status: version 1 Creation date: 03/05/2019

Project co-funded by the European Commission within the H2020 Programme		
Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including the Commission)	
RE	Restricted to a group specified by the consortium (including the	
CO	Confidential, only for members of the consortium (including the	

This project is funded by
the Horizon 2020 Framework
Programme of the European Union

Deliverable D8.3

Guidance on the use of the new strategy for tiered testing and assessment

EuroMix handbook for mixture risk assessment

Main contributors:
Johanna Zilliacus, Anna Beronius, Annika Hanberg, KI
Mirjam Luijten, Jacob van Klaveren, RIVM
Hilko van der Voet, DLO

With input from all EuroMix partners

Deliverable 8.3 “Guidance on the use of the new strategy for tiered testing and assessment” is named “EuroMix handbook for mixture risk assessment” and comprises the main text of the handbook as well as annexes.

May 2019

EuroMix handbook for mixture risk assessment

Johanna Zilliacus, Anna Beronius, Annika Hanberg, Karolinska Institutet, Sweden
Mirjam Luijten and Jacob van Klaveren, RIVM, The Netherlands
Hilko van der Voet, Wageningen University & Research, Biometris, The Netherlands

Table of contents

1 Abbreviations 6
2 Introduction 7
3 EuroMix toolbox 8
4 Problem formulation 9
5 Hazard assessment 12
5.1 Identification and assessment of AOP networks 12
5.2 Collection and assessment of toxicity data from literature 13
5.3 Tiered testing strategies 14
5.4 Grouping of substances based on toxicological considerations 15
5.5 Relative potency factors 17
5.5.1 Application of benchmark dose method to derive RPFs 17
5.5.2 Index substance 19
5.5.3 Use of NOAELs and LOAELs for RPFs 19
5.5.4 Selection of POD 19
5.5.5 In vitro to in vivo (IVIVE) extrapolation of PODs 20
5.5.6 Imputation of missing POD data 20
5.6 Mixture experiments 21
6 Exposure assessment 22
6.1 Probabilistic dietary exposure assessment 22
6.2 Absence of measured concentration data 23
6.3 Non-dietary exposure assessment 23
7 Risk characterisation 24
7.1 Dose addition 24
7.2 Margin of exposure 24
7.3 Selection of main mixtures based on exposure and hazard data 26
8 Tiering approaches 26
9 Uncertainty analysis 26
10 References 28
11 Glossary 32
12 Annexes A - Detailed methodology and templates 34
12.1 Annex A1 - Template for analysis plan 34
12.2 Annex A2 - Organisation of information on AOP networks and toxicity data in the EuroMix toolbox 36
12.3 Annex A3 - Template for extraction of toxicity data for mixture risk assessment 38
12.4 Annex A4 - Template for describing a tiered testing strategy for mixture risk assessment based on AOP networks 40
12.5 Annex A5 - Template for organising data for grouping 41

12.6 Annex A6 - Methodology to derive and use probabilities of assessment group membership 42
12.7 Annex A7 - Methodology to design and analyse mixture experiments 43
12.8 Annex A8 - Possibilities for different tiers using EuroMix methodology and tools 45
12.9 Annex A9 - Template for uncertainty analysis 46
12.10 Annex A10 - Uncertainty modelling in the EuroMix toolbox 47
13 Annexes B - Examples 49
13.1 Annex B1- Example of an analysis plan for mixture risk assessment 49
13.2 Annex B2 - Example on development and assessment of an AOP network for mixture risk assessment 52
13.3 Annex B3 - Example of a tiered testing strategy for mixture risk assessment based on AOP networks 56
13.4 Annex B4 - Example of grouping substances into assessment groups 65
13.5 Annex B5 - Example of calculation of RPFs from in vitro data using benchmark dose method 70
13.6 Annex B6 - Example of imputation of POD using Munro collection of TTC 72
13.7 Annex B7 - Example of risk characterization using dose addition 73
13.8 Annex B8 - Example of uncertainty analysis 74
14 Annexes C - Training material for the EuroMix toolbox 76
14.1 Annex C1 - Training material describing how to group substances based on *in silico* data in the EuroMix toolbox 76
14.2 Annex C2 - Training material for calculation of relative potency factors using the module Relative potency factors in the EuroMix toolbox 76
14.3 Annex C3 - Training material for imputation of missing POD data in the EuroMix toolbox 76
14.4 Annex C4 - Training material for performing a probabilistic dietary exposure assessment of mixtures in the EuroMix toolbox 76
14.5 Annex C5 - Training material for performing exposure assessment using a combination of dietary and non-dietary sources in the EuroMix toolbox 76

1 Abbreviations

3D	three-dimensional
3R	replacement, reduction and refinement
ADI	acceptable daily intake
AO	adverse outcome
AOP	adverse outcome pathway
BMD	benchmark dose
BMR	benchmark response
HBGV	health-based guidance value
IMOE	individual margin of exposure
IPRA	Integrated Probabilistic Risk Assessment
IVIVE	*in vitro* to *in vivo* extrapolation
KE	key event
KER	key event relationship
LOAEL	lowest observed adverse effect level
MIE	molecular initiating event
MOE	margin of exposure
MRL	maximum residue limit
NOAEL	no observed adverse effect level
OHAT	National Toxicology Programme Office of Health Assessment and Translation
PBK	physiologically based kinetic model
POD	point of departure
QSAR	quantitative structure activity relationship
RPF	relative potency factor
SciRAP	Science in Risk Assessment and Policy
SNMU	sparse nonnegative matrix underapproximation
TDI	tolerable daily intake
TTC	threshold of toxicological concern

2 Introduction

The EuroMix handbook for mixture risk assessment describes the methodology and tools for assessing risks of combined exposures to multiple substances developed by the EuroMix project. In the handbook the term “mixture risk assessment” is used as synonym for “risk assessment of combined exposure to multiple substances” for ease of readability. The handbook is consistent with and expands upon the recent documents on mixture risk assessment published by OECD and EFSA (OECD 2018a, EFSA 2019a). The aim is to provide a practical handbook for harmonised application of the EuroMix outcome under consideration of EFSA and OECD guidance and not to repeat the basic principles and information provided in the documents from OECD and EFSA.

The handbook contains concise descriptions of the methodology and tools developed in the EuroMix project. Annexes in the handbook provide detailed information, useful templates and illustrative examples. The mixture risk assessment can be performed using the web-based EuroMix toolbox. The handbook refers to the toolbox but the aim of the handbook is not to provide a step-by-step manual for the toolbox. Detailed information on the EuroMix toolbox is available in the toolbox manual (MCRA 9 2019). Training material for some applications are included as annexes in the handbook.

The EuroMix methodology focuses on component-based mixture risk assessment where substances are grouped based on toxicological considerations. Toxicity and exposure information for each substance in the assessment group is used for estimation of the combined risk using the dose-addition hypothesis and relative potency factors (RPFs). The exposure assessment of mixtures is based on probabilistic methodology considering the individual consumption and concentration data allowing estimation of different percentiles of exposure to the mixture. The focus is on dietary exposure but other exposure routes are also described.

The EuroMix methodology is very flexible, enabling assessment of both data-rich and data-poor substances. The handbook and EuroMix toolbox can also be applied for substances grouped based on other than toxicological considerations, e.g. structure or exposure considerations.

The handbook starts by introducing the EuroMix toolbox. Thereafter, the handbook describes the EuroMix methodology and tools for the key elements in the framework for mixture risk assessment: problem formulation, hazard assessment, exposure assessment and risk characterisation. Finally, general issues of tiering approaches and uncertainty analysis are described. (Figure 1).

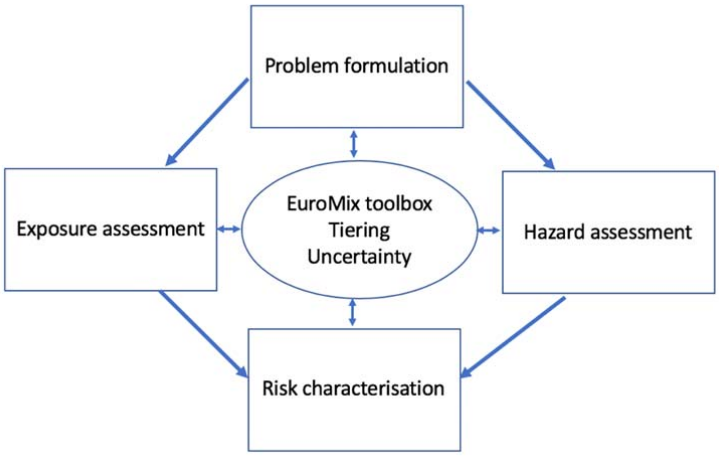


Figure 1. The handbook describes the EuroMix methodology and tools for key elements in the framework for mixture risk assessment, i.e. problem formulation, hazard assessment, exposure assessment and risk characterisation. Furthermore, the EuroMix toolbox, tiering approaches and uncertainty analysis are described.

3 EuroMix toolbox

The EuroMix toolbox, also referred to as MCRA 9, is a web-based toolbox for mixture risk assessment developed in the EuroMix project. It provides a range of tools for application in hazard and exposure assessment of data-rich, as well as data-poor substances. Exposure and toxicity data can be uploaded and used for calculation of e.g. exposure levels, RPFs and risk levels.

The data and models of the toolbox are organized in modules. Each module represents a certain type of data, which can be computed from data provided by other (sub)modules, or the data may be uploaded directly into the toolbox. For each module, an action can be created to configure and run the module. When running an action in the toolbox, the module produces output of its associated data type (which can be used as input for other modules), and a report is generated of the selected data, the model selection and settings, and the module and all intermediate (i.e. sub-modules) results. Figure 2 shows the modules in the toolbox. This handbook will refer to the relevant modules in the toolbox that can be used for the specific applications. Detailed information on the EuroMix toolbox is available in the toolbox manual (MCRA 9 2019).

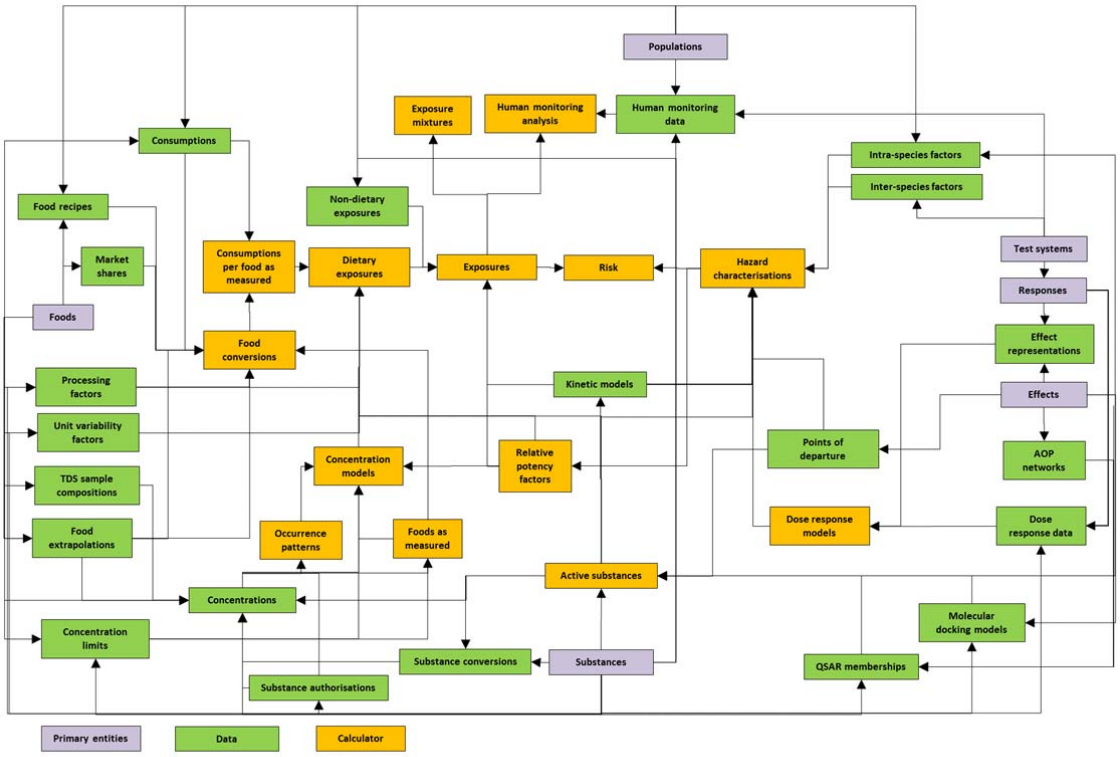


Figure 2. The figure shows the modules in the EuroMix toolbox. The handbook refers to the modules when applicable. Not all modules and relations are fully shown in the figure, for a full description and details please see https://mcra-test.rivm.nl/EuroMix/WebApp/manual/index.html

Annexes

Annex C1-5 contains training material describing how to use some of the modules in the EuroMix toolbox.

4 Problem formulation

Problem formulation is a systematic and often iterative process of defining the purpose and scope of a risk assessment, such as appropriate population groups to be evaluated, relevant substances to be considered, the regulatory goal and intended outcome/use of the assessment, as well as the boundaries of the analysis. The outcome of the problem formulation is an analysis plan, which provides details for how the assessment will be carried out. The problem formulation is often based on a request from risk managers and is developed in dialogue between the risk managers and risk assessors. Several organisations have recently discussed and provided guidance for problem formulation in mixture risk assessment (ATSDR 2018; EFSA 2019a; Meek 2011; OECD 2018a; Solomon 2016). The methodology for the problem formulation described in the handbook is based on this work and adapted to the EuroMix methodology, i.e. a component-based mixture risk assessment, using toxicological considerations as the basis for grouping substances into relevant assessment groups. The methodology for the problem formulation can also be adapted and applied for assessments using grouping based on other than toxicological considerations.

The problem formulation should include consideration of the following steps that are described below: Risk assessment question, Description of the mixture, Conceptual model, Methodological approach and Analysis plan.

Risk assessment question

The risk assessment question is often received from the risk managers. It may include descriptors for the (sub-)population for which the assessment is conducted, the type of substances that are of interest, the sources and/or routes of exposure, as well as the type of effect that is being assessed. The initial risk assessment question does not have to be very detailed and can be refined as needed as the assessment progresses.

Description of the mixture

The aim of this step is to identify whether a mixture risk assessment is required. Description of the mixture should firstly include a description of the concern in regard to the adverse health effect in the population i.e. that there is sufficient evidence (or suspicion) for a common target organ, effect or mode of action, that is of relevance for human health. Secondly, it should address if there is sufficient evidence of co-exposure to the substances identified in the assessment group. It should be noted that co-exposure refers to the internal exposure to the substances. In other words, there can be co-exposure even if external exposure does not occur at the same time. Toxicokinetics of the compounds will also influence the potential for co-exposure. This step requires (preliminary) investigation of available data for both exposure and toxicity. Based on the outcome of this step the decision is made to perform a mixture risk assessment or not.

Conceptual model

The conceptual model aims to define and describe the boundaries for answering the risk assessment question. It provides the basis and rationale for the methodologies applied in the assessment. The conceptual model should describe:

- The regulatory framework and remit under which the assessment is being conducted.
- Substance categories included.
- The relevant exposure sources and routes.
- The appropriate population (sub)group(s), e.g. described by sex, age, occupation, country.
- The toxicological effect being assessed and the level for grouping, i.e. common target organ, common effect/adverse outcome (AO) or common specific mode of action/adverse outcome pathway (AOP).

Methodological approach

In this step, the following aspects related to the applied methodology should be explicitly considered and reported.

- The data availability for both hazard and exposure assessment (“data rich” vs “data poor”) should be described in general terms, including the type of data.
- Whether AOP networks will be used to support grouping of substances into assessment groups and/or identification of measurable effect/key events. (see section 5.1).
- The method and considerations for collection of toxicity data from literature (see sections 5.2).
- The considerations for generation of toxicity data (see section 5.3).
- Methodology and data used for grouping of substances based on toxicological considerations (see section 5.4).
- The calculation of RPFs, choice of index substance and selection of point of departure (POD) (see section 5.5).
- The method for extrapolating between *in vitro* and *in vivo* studies (see section 5.5.5).
- The approach for dealing with lack of toxicity data (see section 5.5.6).
- The methods and considerations for collection of consumption data (see section 6.1).

- The methods and considerations for collection of concentration data (see section 6.1)
- The approach for dealing with non-detects, i.e. concentration measurements below the limit of detection (see section 6.1)
- The approach for dealing with lack of concentration data (see section 6.2).
- The approach for dealing with conversion of food-as-eaten, as found in consumption data, to food-as-measured and for specifying processing factors, for the expected change in concentration due to food processing step such as cooking or peeling (see section 6.1).
- Use of acute or chronic model for exposure assessment (see section 6.1).
- The choice of model for exposure assessment, i.e. whether deterministic or probabilistic (see section 6.1).
- Whether a non-dietary exposure is to be included and if so which external model will generate the exposures (see section 6.3).
- The rationale for deviating from the assumption of dose addition, i.e. in cases where response addition or interactions (synergism or antagonism) are considered more appropriate (see section 7.1).
- The choice of risk metrics to be used (e.g. Margin of exposure (MOE) and different percentiles of the exposure distribution to be used (see section 7.2).

Tiering of the mixture risk assessment refers to the possibility of performing different steps of the assessment using simple, conservative approaches at lower tiers and more advanced approaches requiring more data at higher tiers. If a conservative lower tier assessment indicates that the MOE is sufficiently protective, the assessment does not have to be refined and proceed to a higher tier (see section 8). The problem formulation should specify:

- Which initial tiers should be used for the different parts (hazard assessment, exposure assessment, risk characterisation) of the assessment.
- The methodology to identify risk drivers in a lower tier assessment.
- Which higher tiers are planned to be used in cases when lower tier assessment does not provide sufficient protection.

A description of the method for uncertainty analysis and planned modelling of the uncertainty should be included in the problem formulation (see section 9).

An estimation of the type of expertise, number of persons and other resources needed as well as an estimated timeframe should be described. Plan for stakeholder consultation and peer review should be included.

Analysis plan

The outcome of the steps above is summarised in an analysis plan that describes the planned mixture risk assessment. The problem formulation should be viewed as an iterative process and refinements of the analysis plan may become necessary as the risk assessment progresses and information is gathered.

Annexes

Annex A1 provides a template for the analysis plan.
Annex B1 provides an example of an analysis plan for mixture risk assessment.

5 Hazard assessment

The hazard assessment includes grouping of substances into assessment groups and quantification of the relative potency of the included substances (EFSA 2019a, OECD 2018a). The EuroMix methodology focuses on grouping based on toxicological considerations, but some of the methods and tools can also be applied for substances grouped based on other than toxicological considerations, e.g. structure or exposure. The EuroMix methodology is based on the dose-addition hypothesis using RPFs. The method is flexible and the RPFs can be based on acceptable daily intake (ADI)/tolerable daily intake (TDI), no observed adverse effect level (NOAEL)/benchmark dose (BMD) for the critical effect of the substance that is the basis for setting the ADI/ TDI or the NOAEL/BMD for the specific effect that is the focus of the mixture risk assessment. This section describes methodology and tools for using AOP networks in hazard assessment of mixtures, identification and generation of toxicity data, grouping of substances, quantification of the relative potency and testing of mixtures.

5.1 Identification and assessment of AOP networks

The concept of AOP networks can be useful for mixture risk assessment to support grouping of substances into assessment groups and to identify upstream KEs that can provide toxicity data for RPFs (EFSA 2019a, OECD 2018a). However, mixture risk assessment in the EuroMix toolbox does not require any information from AOPs, only one effect has to be specified for the assessment.

An AOP describes the pathway from a molecular initiating event (MIE), i.e. the interaction between the substance and biological target (e.g. receptor, enzyme), via subsequent steps at molecular, cellular, tissue and organ levels (key events, KEs) to the adverse outcome (AO) in an individual. Multiple AOPs can form an AOP network by converging at the same AO, and/or sharing MIEs or other KEs. The methodology to develop and assess AOPs is described in detail in the OECD Users' handbook supplement to the guidance document for developing and assessing AOPs (OECD 2018b).

The following methodology can be used to identify and assess AOP networks for use in mixture risk assessment.

- First, any existing AOPs for the AO published in the AOP wiki (https://aopwiki.org) or literature should be identified and used as basis for any further development. In case none are available, the development of a new AOP can start identifying the AO and thereafter identifying KEs leading to the AO.
- KEs leading to the AO are identified by searching the literature for evidence linking the KEs to each other and to the AO, using the methodology described in OECD 2018b. It is most useful to identify KEs that can be easily measured to inform grouping and provide toxicity data for RPFs.
- It is not necessary to develop a complete AOP. Even an AOP with only a single KE in addition to the AO may be useful.
- When the AOP has been postulated, it should be assessed, as described in the OECD AOP handbook (OECD 2018b). The assessment includes evaluation of the biological plausibility and empirical support for the KE relationships linking the KEs, as well as evidence supporting the essentiality of the KEs.
- The AOP network, including the MIEs, KEs, AOs and KE relationships should be described in tables for use in the EuroMix toolbox (Annex A2).

EuroMix toolbox

The modules AOP networks and Effects in the EuroMix toolbox are used for describing the AOP network and the MIEs, KEs and AOs (effects) in the AOP network.

Annexes

Annex A2 details how AOP networks, KEs and KE relationships are described for use in the EuroMix toolbox.

Annex B2 includes an example of development and assessment of an AOP network for mixture risk assessment.

5.2 Collection and assessment of toxicity data from literature

Toxicity data is needed in mixture risk assessment for grouping of substances into assessment groups and for calculation of RPFs for substances in the assessment group. Toxicity data can be generated specifically for the assessment at hand using a tiered testing strategy, see section 5.3. However, in many cases there is a need to identify, collect and assess toxicity data from the available literature, e.g. from dossiers, reports or scientific publications. Systematic and transparent methods are important to ensure that the data used for the mixture risk assessment is relevant and reliable. Principles from systematic review and weight of evidence methodologies are useful (EFSA 2010, EFSA 2011, EFSA 2017a). AOP networks can be used to identify the effects that should be included in the data collection and form the basis for the search strategy. The methodology described here is based on previously used methodology for collecting toxicity data for mixture risk assessment (RIVM, ICPS, ANSES 2013, 2016) and EuroMix deliverable 2.1 (EuroMix 2016a).

The following methodology can be used to collect and assess toxicity data from literature for mixture risk assessment.

- Purpose of data collection
  The problem formulation at the start of the mixture risk assessment will identify the purpose of the data collection, e.g., grouping of substances into assessment groups and/or identification of toxicity data to be used for calculation of RPFs.
- Search for studies from reports, dossiers and scientific publications
  The first step is to identify the sources where relevant studies can be found, such as PubMed for scientific publications or EFSA databases. Thereafter, a search strategy including e.g. search terms for PubMed searches is identified and the search is performed. If possible, the search should be carried out by experts, such as information specialists and preferably in two or more databases.
- Select the studies that contain relevant data
  When the literature search has been completed, relevant studies are selected based on specific selection criteria, related e.g. to the identity to the substances, effects or study types.
- Collect data from the studies
  The next step is to extract the important data from the studies into an Excel file. The data to be extracted will partly depend on the purpose of the data collection and the specific effects or type of studies that are addressed. Annex A3 lists data to be extracted from *in vivo* and *in vitro* studies for the purpose of either grouping or RPFs. The template is also available as an Excel file including drop-down lists that facilitate data extraction.
- Assess the data for reliability and relevance
  The reliability (quality) of the data must be evaluated. Tools for data evaluation such as the SciRAP tool (Beronius 2018, www.scirap.org) or OHAT risk of bias tool (NTP 2015) can be used. For example, data from well reported research studies of high methodological quality or from studies based on OECD test guidelines might be considered as more reliable than poorly reported or poorly performed studies. The relevance of the data for the specific purpose should also be evaluated. For example, data from *in vitro* studies for upstream KEs might be considered as less relevant than *in vivo* studies measuring the AO. Scoring systems for reliability and relevance can be used in case it facilitates the assessment and can be reported in the data extraction Excel file.

It is recommended that the selection of studies, extraction of data and assessment of data is done by two reviewers and that any disagreements between them are resolved by discussion or involving a third reviewer.

Annexes

Annex A3 provides a template for extraction of toxicity data from *in vivo* and *in vitro* studies for the purpose of either grouping or RPFs. The table follows the data formats used in the EuroMix toolbox. The template is also available as an Excel file including drop-down lists that facilitate data extraction.

5.3 Tiered testing strategies

One of the challenges of mixture risk assessment is availability of toxicity data for all substances that are included in the assessment group. In case there is no existing data it may become necessary to generate data specifically for the assessment at hand. A tiered testing strategy can based on data from *in silico* modelling set priorities for *in vitro* testing that further informs and sets priorities for *in vivo* testing (OECD 2018a). The use of *in silico* and *in vitro* data to the extent possible supports the 3R principles (replacement, reduction and refinement) and avoids animal testing.

In silico data from quantitative structure activity relationship (QSAR) models can predict toxicity at organ level (e.g. hepatotoxicity) or at the level of an effect/AO (e.g. liver steatosis). QSAR models can also predict activation of MIEs, such as nuclear receptor activation. The QSAR data can be used for grouping of substances and for prioritisation of testing. *In silico* data from molecular docking, using either experimental three-dimensional (3D) structures, when available, or comparative 3D models, can be used to estimate binding energies to receptors and enzymes and thereby provide low tier toxicity data for RPFs (EuroMix 2016b). *In vitro* data from e.g. cell lines, organ cultures or zebrafish embryos can be useful for grouping, to derive potency information and for prioritisation of further testing *in vivo* (EuroMix 2017b, 2018a, Luckert 2018).

AOP networks can provide the basis for planning strategic testing at different levels of biological organisation (OECD 2016). The AOP network makes it possible to identify effects/KEs that can be tested using *in silico* models, *in vitro* and *in vivo* assays.

The following methodology can be used to develop a tiered testing strategy for mixture risk assessment based on AOP networks.

- Identification of the KEs in the AOP network that can provide information for grouping and/or RPFs.
- Identification of *in silico* models (e.g. QSAR or molecular docking models), *in vitro* and *in vivo* assays for measurement the KEs.
- Assessment and description of the relevance of the *in silico* models and *in vitro* and *in vivo* assays used for measuring of the KEs. The assessment should take into consideration e.g. the applicability domain of the *in silico* model or the relevance of the specific measured response and the *in vitro* and *in vivo* test system.
- Assessment and description of the reliability of the outcome from the *in silico* models and *in vitro* and *in vivo* assays. The assessment should take into consideration e.g. accuracy, sensitivity and specificity of the *in silico* model and *in vitro* and *in vivo* assay.
- Assessment of the availability and feasibility, in terms of costs and other resources, for the *in silico* models and *in vitro* and *in vivo* assays.
- Assessment and description of the information provided by the *in silico* models and *in vitro* and *in vivo* assays to support the mixture risk assessment, i.e. for grouping, RPFs and/or prioritisation for further testing.

- Selection of the final *in silico* models and *in vitro* and *in vivo* assays to be included in the tiered testing strategy based on the assessments above.
- The tiered testing strategy can include recommendations on a step-wise approach for the testing, e.g. models/assays to perform first and how to proceed dependent on positive or negative results in the previous model/assay.
- Description of the selected *in silico* models and *in vitro* and *in vivo* assays (i.e. test systems and responses) in tables for use in the EuroMix toolbox.

Toxicity data from *in silico*, *in vitro* and *in vivo* studies can be used in the EuroMix toolbox. The data is organised based on the concept of AOP networks and each *in silico* model, *in vitro* assay or *in vivo* test is matched to a specific KE. The organisation of the toxicity data for the EuroMix toolbox is described in Annex A2.

EuroMix toolbox

The modules Test systems and Responses in the EuroMix toolbox are used to identify *in vitro* and *in vivo* assays. The module Effect Representations connects the assays to the KEs in the AOP network. The module Dose response data contains dose response data from *in vitro* and *in vivo* studies. The module Dose response models can be used to derive BMDs from these data. Externally specified values for BMDs and/or NOAELs can be specified in the module Points of departure. The modules QSAR membership models and Molecular docking models are used to specify *in silico* data.

Annexes

Annex A2 details how *in vitro* and *in vivo* assays and dose response data from these assays are described for use in the EuroMix toolbox.
Annex A4 provides a template for describing a tiered testing strategy for mixture risk assessment based on AOP networks.
Annex B3 includes an example of a tiered testing strategy for mixture risk assessment based on AOP networks.

5.4 Grouping of substances based on toxicological considerations

Substances can be grouped into relevant assessment groups based on different approaches. EuroMix provides methodology for grouping based on toxicological considerations. Dose addition is the default recommended model for mixture risk assessment (EFSA 2019a, OECD 2018a) and is used in the EuroMix methodology for substances grouped into the same assessment group.

AOP networks can provide a framework for grouping of substances. Grouping based on toxicity can be performed at different levels of biological organisation, i.e. common target organ, common effect/AO or common specific mode of action/AOP. Grouping at the level of a common target organ may be necessary for some data-poor substances for which no information on specific effects/AO in the target organ is available. Grouping at the level of common effect/AO will probably be used in most cases. Since several AOPs can form an AOP network converging to the same AO, substances that act via any of these AOPs can be grouped together at the level of common effect/AO. In certain cases, evidence may indicate that the substances cause the AO via separate independent AOPs and the model for dose addition does not appropriately describe the combined effect of the separate AOPs. In such cases, the substances would be grouped based on the specific mode of action/AOP and the model for response addition could potentially be used.

The decision whether a substance should or should not belong to an assessment group should be made based on all available relevant evidence. However, in many cases it is uncertain which substances should belong to a specific assessment group. The uncertainty about group membership

can be expressed as a probability of group membership and used in modelling in the EuroMix toolbox as described in Annex A6.

Specific criteria related to exposure and toxicity can also be used to decide on assessment group membership, such as exclusion of substances below a specified exposure level or exclusion of substances for which the POD of the critical effect, that is the basis for setting the ADI/TDI, is lower than the POD of the specific effect that is the focus of the mixture risk assessment. Such criteria should be described and justified in the problem formulation.

The following methodology can be applied for grouping at any level, i.e. common target organ, common effect/AO or common specific mode of action/AOP, and irrespective of whether the group membership is expressed as a probability or as "included/not included". The described methodology is based on EuroMix deliverable 2.1 (EuroMix 2016a) and consistent with and expands upon methodology in EFSA 2018b, 2019b.

- Level of grouping
  The level of grouping, common target organ, common effect/AO or common specific mode of action/AOP, is first decided.
- AOP network
  An AOP network for the AO is identified, when needed. In cases when grouping is at the level of common effect/AO and toxicity data is available for the AO for all substances in the assessment group, information on the AOP network might not necessary to decide on grouping. However, in cases where toxicity data on the AO is missing for some or all substances, toxicity data for KEs in the AOP network can be used to inform the grouping.
- Substance category to be assessed
  Substance category to be assessed is identified in the problem formulation, e.g. pesticides approved in Europe or contaminants identified by human biomonitoring.
- Collection of data
  Toxicological data for the substances is collected from the literature and relevant databases. The data can be from *in silico*, *in vitro*, *in vivo* or human studies and can be related to the AO or any KEs in the AOP network. In case data from *in vivo* studies for the AO is available, additional data might not be needed. Data collection can be done in a tiered manner, and additional data is only required when the uncertainty of group membership is high. In the special case only *in silico* data is available, grouping can be done based on the results from the *in silico* models only. See also sections 5.2 and 5.3.
- Organising the data
  The data is organised into lines of evidence. For example, data can be arranged for each KE and for the AO and can be further organised according to data from *in silico*, *in vitro*, *in vivo* or human studies. A template is provided in Annex A5.
- Assessing the data for reliability and relevance
  The reliability (quality) of the data is evaluated. See section 5.2. The relevance of the data for grouping into assessment groups is also evaluated. For example, data from *in silico* and from *in vitro* studies for up-stream KEs might be considered as less relevant than *in vivo* studies measuring the AO. Scoring systems for reliability and relevance can be used in case it facilitates the assessment and the following steps.
- Decision of group membership
  The data is assessed considering the reliability and relevance in a weight of evidence approach. Well-organised data, including information and justification for evaluation of reliability and relevance, facilitates the decision-making. The decision on group membership should be done by at least two experts. Processes should be in place to resolve any disagreements between experts.

Formal expert knowledge elicitation can be used, and is preferable when quantifying probabilities of group membership (EFSA 2014, 2018b, 2019b).

- Reporting of group membership
  The group membership for each substance is expressed as either 0 (not included) or 1 (included) or as a value between 0-1 indicating the probability for belonging to the assessment group. The group membership is reported in a table for use in the EuroMix toolbox.

EuroMix toolbox

The module Active substances in the EuroMix toolbox is used to identify the substances included in the assessment group.

Annexes

Annex A5 provides a template for organising data for grouping.
Annex A6 describes methodology to derive and use probabilities of assessment group membership
Annex B4 provides an example of grouping substances into assessment groups.
Annex C1 provides training material describing how to group substances based on *in silico* data in the EuroMix toolbox.

5.5 Relative potency factors

Dose addition using RPFs is the primary method for modelling the risk of mixtures in the EuroMix toolbox. The exposure of each substance is multiplied with the RPF of the substance and the potency-scaled exposures are summed. The RPF of each specific substance is derived by dividing the POD of the index substance with the POD of the specific substance. The POD value can be a benchmark dose (BMD) from benchmark dose modelling or a No Observed Adverse Effect Level (NOAEL). The method is flexible, the RPFs can be based on the POD for the critical effect of the substance that is the basis for setting the ADI/TDI or the POD for the specific effect that is the focus of the mixture risk assessment. ADI/TDI can be used instead of POD, i.e. the POD divided by the assessment factors, which should be taken into consideration when interpreting the MOE, see section 7.2. The choice of using POD from a critical or specific effect or a ADI/TDI instead of POD depends on the tier of the mixture risk assessment and data availability.

The equations related to RPFs are described here:
Toxicity of index substance: POD_{index}
Toxicity of each substance in mixture: POD_1, POD_2, POD_3, POD_4 …
Relative potency factor (RPF) for each substance: $RPF_1 = POD_{index} / POD_1$
Exposure to mixture is scaled based on RPFs: $Exp_{mix} = Exp_1 \times RPF_1 + Exp_2 \times RPF_2 + \ldots$

5.5.1 Application of benchmark dose method to derive RPFs

Historically the NOAEL has been used as the POD. More recently, the benchmark dose method has been developed as a more scientific approach for dose response assessment and estimation of a POD (Slob 2002, US EPA 2012, EFSA 2017b). When using the benchmark dose method, a specific and appropriate effect size level (benchmark response, BMR) needs to be chosen for estimation of the corresponding BMD. The BMR is usually expressed as a relative deviation from the mean response in the control group. The basis for POD estimations is usually *in vivo* data on AOs. According to the EFSA guidance (2017b), the default BMR for continuous *in vivo* data is 5% (increase or decrease compared to the background/control level, see Figure 2). For quantal *in vivo* data the recommended BMR is 10%. These defaults/recommendations of BMR levels are meant for cases where the aim is to derive a POD and subsequently a health-based guidance value (HBGV). In addition, EFSA states that the benchmark dose method can also be applied on other types of data (like *in vitro* data) and for other purposes (like estimating RPFs). It is also mentioned that in certain cases the selected BMR level may deviate from these defaults, if there are biological or statistical reasons for that.

The benchmark dose method can be used to derive RPFs in two different scenarios. In the first scenario the purpose is to derive BMDs to calculate RPFs for a group of substances, but the BMDs will not be used as a POD for the risk assessment. In this case it is not necessary to select a BMR that reflects a no effect level, but a higher BMR can be chosen that is statistically more robust, and anywhere between the minimal and maximal observed responses, provided the dose-response curves are parallel. In the mixture risk assessment, the potency-scaled exposure derived from these RPFs should be compared to a BMD for the index substance that is derived using a BMR that reflects the no effect level as described in the EFSA guidance (EFSA 2017b). In the second scenario, BMDs are derived to be used as PODs in a risk assessment, and not only to calculate RPFs. In this case the BMR should be chosen to reflect a no effect level, according to the EFSA guidance. The recommendation in the EFSA guidance is for *in vivo* data. For *in vitro* data a higher BMR might be needed to reflect the no effect level. The choice of BMR for *in vitro* data has to be done case by case and the rationale for the selected BMR should be described. Use of a BMD based on *in vitro data* also requires *in vitro* to *in vivo* extrapolation, see section 5.5.5.

The benchmark dose software Proast is integrated in the EuroMix toolbox allowing benchmark dose analysis of dose response data from *in vitro* and *in vivo* studies to be done within the toolbox (https://www.rivm.nl/en/proast).

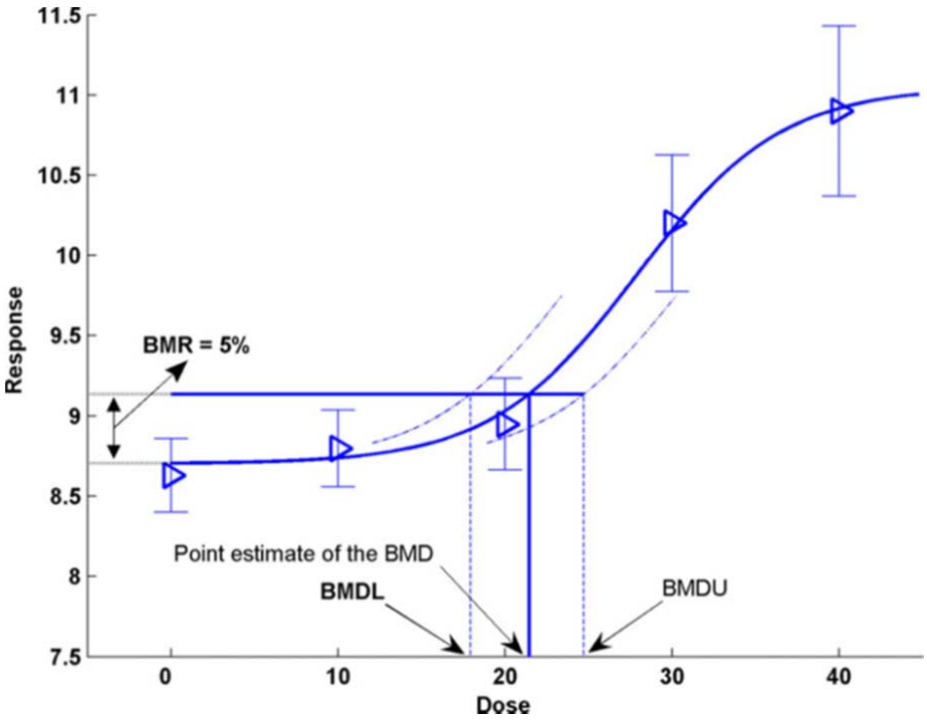


Figure 2. Benchmark dose modelling of continuous dose-response data (from EFSA 2017b).

5.5.2 Index substance

The index substance is important in the RPF approach and should be chosen considering the following criteria:

- confidence that the substance is representative for the specific assessment group
- confidence that the substance causes the effect that is the basis for the risk assessment
- the POD is derived from an *in vivo* study for the effect in focus for the mixture risk assessment
- quality and quantity of toxicity data, resulting in a high confidence in the POD

It should be noted that, the index substance does not have to be the most toxic substance (i.e. lowest POD) in the assessment group.

For estimation of the BMD of the index substance, the biological relevance of the BMR-level is critical as the BMD is used for estimating the POD for the mixture risk assessment and should relate to an effect size that is considered a no effect level. As the POD for the mixture risk assessment should be based on the adverse effect, and take into consideration also toxicokinetics and toxicodynamics in an intact organism, it is current practice that the POD of the index substance is based on *in vivo* data on the AO. In the future, it is foreseen that in cases where an *in vivo* study is not available other data can be used. If the POD is based on *in vitro* data, these will need to be extrapolated to *in vivo* using physiologically based kinetic (PBK) and *in vitro* to *in vivo* extrapolation (IVIVE) models in order to estimate a corresponding dose/exposure level. Currently the precision of PBK models as well as availability of high-confidence AOPs are too limited for *in vitro* data being used as the basis for deriving a POD for the index substance.

5.5.3 Use of NOAELs and LOAELs for RPFs

A POD can be in the form of a BMD from Benchmark dose modelling or a NOAEL or lowest observed adverse effect level (LOAEL). Whenever possible, a BMD should be used. In case a study only reports a NOAEL, it should be considered if it is possible to extract the dose-response data and calculate a BMD using benchmark dose modelling. In case different types of PODs are used for different substances in the assessment group there is a need for conversion. In the EuroMix toolbox conversion factors can be used, based on the WHO Guidance document on evaluating and expressing uncertainty in hazard characterization (WHO 2018).

5.5.4 Selection of POD

There might be several different PODs available from different studies measuring the same response or different responses. The selection of POD to be used for the mixture risk assessment should consider the following.

- Comparability within the assessment group
  The selected PODs for the substances within the assessment group should be comparable. Therefore, the PODs should be for the same response for all substances, i.e. the same outcome measured using the same study design. In cases where this is not possible, similar responses should be selected.

- Responses from different KEs in the AOP network
  PODs might be available for all substances in the assessment group from different responses measuring either upstream or downstream KEs in the AOP network. In these cases, the relevance of the responses, including the study design, for the mixture risk assessment should be considered. Responses measuring KEs close to the AO are probably more relevant than responses measuring the MIEs or upstream KEs.

- Several PODs for same response
  In the case that several PODs are available for a substance for the same response but from different studies, either the most reliable or the most conservative POD can be selected. An overall POD can also be chosen by considering the studies together and choosing the highest NOAEL that provides a reasonable margin to the lowest LOAEL (IPCS 2009).

- Selection of POD in the EuroMix toolbox
  The EuroMix toolbox also allows for automatic selection of the POD for a substance in case several PODs have been uploaded into the toolbox. The lowest POD (conservative) or the mean POD can be chosen.

5.5.5 In vitro to in vivo (IVIVE) extrapolation of PODs

Traditionally, toxicity data is expressed as external oral dose from an *in vivo* study and can easily be compared to the dietary exposure levels. In mixture risk assessment toxicity data for some (or all) substances might be based on *in vitro* studies. In such cases, the *in vitro* derived internal POD can be extrapolated to an external POD, considering the kinetics of the substance. Alternatively, the oral dietary exposure can be extrapolated to an internal exposure that can be compared to the *in vitro* POD. Kinetic models can be used for the extrapolations in either direction. In a lower tier, kinetic models are specified as absorption factors, that can be the same for all substances or be substance specific, depending on the available information. In a higher tier, substance specific PBK models can be used to derive organ specific factors. For example, to extrapolate a POD derived from data using liver cells to an external POD. The PBK models can be built on the basis of experimental substance specific data or can be generic and use QSAR derived parameters (EuroMix 2018b). In the EuroMix toolbox IVIVE extrapolations can be performed using either kinetic factors of PBK models.

5.5.6 Imputation of missing POD data

One of the challenges of mixture risk assessment is to identify toxicity data for all substances included in the assessment group. Substances should not be excluded from the assessment due to lack of information on POD. Instead the missing POD data can be imputed. Different approaches how to impute missing PODs developed in EuroMix (EuroMix 2017a) are described here.

- POD based on Munro collection of TTC values
  POD values can be imputed based on the threshold of toxicological concern (TTC) concept (EFSA 2018c). Munro et al 1996 compiled a database of NOAEL values for 613 substances from oral subchronic and chronic studies in rodents and rabbits. The most conservative NOAELs were selected for each substance. The substances were also classified into the three Cramer classes based on a decision tree mainly related to chemical structure (Cramer 1978). The 5^{th} percentile of the NOAELs in each class was calculated and divided by an uncertainty factor of 100 to derive TTC values for the three classes: 30, 9.0, 1.5 µg/kg bw per day for Cramer classes I, II and III, respectively. The appropriateness of these TTC values was confirmed in a recent review organised by EFSA and the WHO (EFSA WHO 2016). The Munro database can be used in mixture risk assessment for substances without any specific POD information. Either the 5^{th} percentile or the mean of the NOAELs can be used. In the EuroMix toolbox, the uncertainty about the right NOAEL value can be included in the modelling by not only using the 5^{th} percentile or mean, but by randomly sampling among the NOAELs in the uncertainty iterations of the Monte Carlo simulation.

- POD based on existing PODs in the assessment group
  The Munro collection of NOAELs is based on different types of endpoints, e.g. specific organ toxicities, reproductive toxicity or body weight changes. Therefore, an alternative for the missing PODs is to use a collection of PODs for the specific effect in the mixture risk assessment. The

assumption is that the effect-specific PODs would more closely reflect the true POD for the substance and that using them is a less conservative but more realistic approach compared to using the Munro NOAELs. PODs for the other substances in the assessment group are collected in the process of the hazard assessment and reported in a table that can be used for either deriving the 5th percentile or the average of the PODs. The uncertainty can also be modelled by random sampling from all PODs.

EuroMix toolbox

The module Relative potency factors in the EuroMix toolbox is used to specify or calculate the RPFs. The module Dose Response Models is used to perform the Benchmark dose modelling using the Proast software that is integrated into the toolbox. The Point of Departure module contains POD data uploaded into the toolbox. The Hazard characterisations module is used to extrapolate between different types of POD, to select one POD if several are available, to impute missing PODs and for IVIVE. The Kinetic models module contains the models needed for IVIVE.

Annexes

Annex B5 provides an example of calculation of RPFs from *in vitro* data using benchmark dose method
Annex B6 provides an example of imputation of POD using Munro collection of TTC
Annex C2 provides training material describing how to use the module Relative potency factors for calculating RPFs in the EuroMix toolbox.
Annex C3 provides training material describing how to impute missing POD data in the EuroMix toolbox.

5.6 Mixture experiments

Interaction between substances in a mixture, due to chemical-chemical, toxicokinetic or toxicodynamic interactions, can potentially result in synergism or antagonism. In such cases, the combined effect will deviate from dose addition. However, available evidence suggests that synergism is rare at dietary exposure levels (EFSA 2019a, OECD 2018a). Testing of mixtures of substances can provide information on deviations from dose addition.

The following methodology can be used to design and analyse mixture experiments, using binary mixtures.

- The mixture experiment should be performed using doses with equal potency. Therefore, relative potency information for the individual substances is required to design the mixture experiment. It is preferable that the individual substances first are tested in the selected assay and the relative potency is calculated. In case this is not possible, relative potency derived from other assays can be used, e.g. relative potency from *in vitro* assays can be used to design *in vivo* testing. The relative potency of the individual substances can be identified using the benchmark dose method (see section 5.5.1).
- The mixture experiment should include several doses of each individual substance and of the mixture of the two substances. The mixture doses should be equipotent to the individual doses (for details see Annex A8).
- The result is analysed using the benchmark dose method in a combined analysis of the individual substances and the mixture (for details see Annex A8).

EuroMix toolbox

The module Relative potency factors in the EuroMix toolbox is used to calculate RPFs and analyse mixture experiments.

Annexes

Annex A7 describes the detailed methodology to design and analyse mixture experiments.

6 Exposure assessment

6.1 Probabilistic dietary exposure assessment

In the EuroMix toolbox dietary exposure is estimated using probabilistic approach (EFSA 2012). Probabilistic exposure assessment can provide a distribution of the exposure estimate. Quantification of uncertainty can also be performed.

The exposure can be modelled based on acute or chronic exposures. The selection should be based on whether the effect that is the focus of the assessment can be caused by a single exposure in a short time (acute) or by repeated exposure over a longer time (chronic).

Probabilistic modelling is based on random selection of one individual (for chronic risks) or individual-day (for acute risks) from the consumption database. Further, for each food item that the individual has consumed, concentration data are randomly selected from available concentration data for that food item (for acute risks) or the mean concentration is used (chronic risks). The consumption and concentration data are multiplied for each food item and added up for all food items the individual has consumed to estimate the individual's total dietary exposure. For acute risks, this is repeated in a large number of iterations, e.g. 100 000 times to provide an effective Monte Carlo integration of the distributions of consumption and concentration. For chronic risks, the exposure distribution is represented by distribution of the consumptions times the mean concentrations summed over food items and averaged over the survey days of each individual (Observed Individual Means method, this is a simple model chosen as default in EFSA 2012, and which can be refined if needed using more advanced models that are also implemented in the toolbox). The relative potency of the substance is also included in the calculation by multiplying the exposure of a substance with its RPF to express the exposure as equivalents of the index substance. The equivalents of the index substance are summed for each individual (chronic) or individual day (acute) to obtain a distribution of the exposure to the mixture.

Data requirements and principles for probabilistic dietary exposure assessment for single substances and for a mixture are similar. Food consumption data from food consumption surveys and concentration data from measurement of levels of substances in food items are needed. Conversion of the foods-as-eaten, as found in the food consumption data, to foods-as-measured, which are the foods for which concentration data exists, using food recipes is also required.

EuroMix toolbox

The EuroMix toolbox contains several modules related to the description of Foods, Consumptions and Concentrations. The module Consumptions is used to specify the consumption survey data. The module Food recipes is used to specify relations between food-as-eaten and foods-as-measured. The module Concentrations is used to specify concentration data, typically food monitoring data, but also field trial data or data obtained from total diet studies could be used. The module Concentration models allows for more specific modelling aspects of the concentration data, e.g. how to handle left-censored observations (non-detects, reported as 'below a limit of reporting'), and how to address substance conversions due to differences between the definitions of active substances and the substances measured by analytical methods in monitoring programmes (residue definitions). The module Processing factors is used to specify factors for the expected change in concentration due to food processing step such as cooking or peeling. The module Unit variability factors is used to specify

models for deriving concentrations in individual food units (e.g. apples) from the measured values as obtained from composite samples (e.g. 12 mixed apples). The module Dietary Exposures is used for the probabilistic modelling considering the consumption and concentration data as well as food and substance conversions, concentration modelling steps, processing factors and unit variability factors.

Annexes

Annex C4 provides training material describing how perform a probabilistic dietary exposure assessment of mixtures in the EuroMix toolbox.

6.2 Absence of measured concentration data

One of the challenges of mixture risk assessment is the availability of concentration data for all substances included in the assessment.

In the case that no or not enough analytical concentration data is available for a particular substance there are several options. A first option is to extrapolate data from other foods, e.g. assigning concentrations in spinach to other leaf vegetables. This option can only be used when food consumption data is organised using a hierarchical food coding system, such as EFSA FoodEx1 coding system. A second option is to use the substance´s legal limits in food. For example, the legal Maximum Residue Limit (MRL) of the pesticide framework can be used to impute missing concentration data.

EuroMix toolbox

The module Food extrapolation can be used to specify whether concentration data on a food for which this is missing may be extrapolated from another food. The module Concentration limits in the EuroMix toolbox contains data on e.g. MRLs.

6.3 Non-dietary exposure assessment

The EuroMix toolbox also provides the possibility to model the risk from exposure to substances that are present both in dietary and non-dietary sources. In general, a non-dietary exposure source may be relevant to only a subset of the population. These sources are implemented as non-dietary surveys that can be limited by individual properties e.g. age ranges and gender. Non-dietary exposure estimates can be generated by external programs, imported into the EuroMix toolbox and combined with dietary exposure estimates. In the framework of the EuroMix project, coupling with the tools Browse, PACEM or RSExpo was tested. Browse can calculate dermal, oral and inhalation exposures. It was developed to calculate exposure to pesticides by bystanders, residents, operators and workers, due to their proximity to the crop-spraying activities (Kennedy 2016, Kennedy 2019, https://secure.fera.defra.gov.uk/browse). PACEM is a tool that was programmed in R, which calculates dermal, oral and inhalation exposures with an individual-based probabilistic method. It was developed to assess consumer exposure to substances in personal care products (Dudzina 2015) and in household products (Garcia-Hidalgo 2018). In the framework of the EuroMix project, a user-friendly interface was developed (Karrer 2019). RSExpo is a Rshinny application using the R software (R Core Team, 2017) that makes it possible to calculate exposure from several sources for general population. A case study was applied to Pyrethroids (Vanacker 2019).
The non-dietary exposures modelled by these programmes can be imported into the EuroMix toolbox and combined with dietary exposure modelling. Browse and PACEM both include the option to export exposure estimates in a file format required by the EuroMix toolbox. Exposure from dermal, oral and inhalation routes might be substance-specific and can either be transferred from external to internal exposure using simple default absorptions or kinetic models. If the external

program generates uncertainties in the form of multiple exposure sets, these can also be input and used to quantify this uncertainty within the overall assessment.

EuroMix toolbox

The module Non-dietary exposures in the EuroMix toolbox is used for specifying the non-dietary exposure data. The module Exposures aggregates dietary and non-dietary exposures. The module Kinetic models can be used to provide conversions from external to internal exposure.

Annexes

Annex C5 provides training material describing how to perform exposure assessment when using a combination of dietary and non-dietary sources in the EuroMix toolbox.

7 Risk characterisation

7.1 Dose addition

Dose addition is the default recommended model for mixture risk assessment (EFSA 2019a, OECD 2018a) and is used in the EuroMix methodology for substances grouped into the same assessment group. In certain cases, evidence may indicate that the substances cause the AO via separate independent AOPs and the model for dose addition does not appropriately describe the combined effect of the separate AOPs. In such cases, the substances would be grouped based on the specific mode of actions/AOPs and the model for response addition could potentially be used.

According to the concept of dose addition, the combined toxicity of substances with similar toxicity can be estimated from the sum of the doses scaled for their relative toxicity. In EuroMix, dose addition is used with the application of RPFs. Calculation of RPFs is described in section 5.5.

The combined exposure to the substances in the assessment group is calculated by multiplying the exposure of each substance with the RPF of the substance, and adding up these potency-scaled exposures. The combined exposure is now expressed as index substance equivalent.

In the EuroMix toolbox, exposures are modelled probabilistically. The potency-scaled combined exposure can be expressed as a distribution and different percentiles of the exposure distribution can be calculated.

EuroMix toolbox

The module Dietary exposures in the EuroMix toolbox calculates the potency-scaled combined dietary exposure using the dose addition model. The module Exposures provides in the simplest case just these dietary exposures, but can be used to aggregate dietary and non-dietary exposures, or to translate external to internal exposures.

7.2 Margin of exposure

Risk characterisation based on dose addition can use different models to compare the exposure to the hazard. All models are based on the same principle and the main difference is the input used, e.g. use of ADI/TDI, POD for critical or specific effect for the hazard, use of point estimates or distributions for the exposure and use of assessment factors (EFSA 2019a, OECD 2018a). The EuroMix

methodology uses the MOE model using RPFs. The methodology is very flexible and can also be used at a low tier based on ADI/TDI values as the hazard index model.

In the MOE model, the potency-scaled combined exposure is expressed as index substance equivalents and compared with the POD of the index substance. The comparison can be done and visualised in different ways using the MOE concept.

The equations related to MOE are described here:
Toxicity of index substance: POD_{index}
Toxicity of each substance in mixture: POD_1, POD_2, POD_3, POD_4 …
Relative potency factor (RPF) for each substance: $RPF_1 = POD_{index} / POD_1$
Exposure to mixture is scaled based on RPFs: $Exp_{mix} = Exp_1 \times RPF_1 + Exp_2 \times RPF_2 + \ldots$
MOE of mixture: $MOE_{mix} = POD_{index} / Exp_{mix}$

The MOE can be expressed for the mean exposure or for different percentiles of the exposure distribution. An MOE of 100 or more is generally considered acceptable. This is based on the standard assessment factor of 100 used to extrapolate from the animals used in the *in vivo* study to the most sensitive human. It should be noted that in certain cases an MOE larger than 100 may be considered necessary.

The EuroMix toolbox can also calculate a more refined risk assessment using probabilistic hazard characterisation. In the most complete version this implements the Integrated Probabilistic Risk Assessment (IPRA) model (van der Voet 2007, van der Voet 2009). In this approach the PODs are modified by a probabilistic modelling of the assessment factors for inter-species extrapolation and intra-species variability. Specifically, the traditional approach to protect the 'sensitive human' by applying an intra-species factor (typically a fixed value of 10, meant to cover both variability between humans and the uncertainty) is replaced by a distribution of values to represent the sensitivities of all individuals in the population. This distribution is by definition centred around a factor 1 (the average sensitivity factor) and could e.g. be specified as a lognormal distribution with a 95^{th} percentile (representing the 'sensitive individual') of between 2 and 10 (this range specifying the uncertainty expressed as a 95% confidence interval).

When using more refined hazard characterisation than just external PODs, the traditional MOE is replaced by probabilistic generalisations of the MOE concept. In the case of the IPRA model, individual MOEs (IMOEs) are calculated that directly compare human hazard characterisation to human exposures. Note that the standard MOE interpretation criterion of 100 should then be modified. In the IPRA case an IMOE > 1 is considered acceptable, because the assessment factors are already included in the IMOE calculation.

Therefore, the MOE concept should be used carefully and it should be clearly reported how the exposure and POD is expressed.

In the EuroMix toolbox the potency-scaled combined exposure and the POD for the index substance are plotted against each other, the exposure on the x-axis and the POD on the y-axis. The uncertainty both in the exposure and the POD can be expressed in such a plot using ellipses. The single substances can also be represented in the same way. Colour coding of the plot can be used, the diagonal indicating MOE=100 is yellow, and areas MOE>100 are green and MOE<100 are red.

EuroMix toolbox

The module Dietary exposures in the EuroMix toolbox calculates the potency-scaled combined exposure using the dose addition model and expresses the MOE. The module Hazard characterisations can be used to select either the simple PODs or modified, potentially probabilistic

versions thereof. The module Risks plots the potency-scaled combined exposures and the POD for the index substance to visualise the results. The module Risks also calculates MOE or probabilistic generalisations thereof.

Annexes

Annex B7 provides an example of risk characterisation using dose addition

7.3 Selection of main mixtures based on exposure and hazard data

The EuroMix toolbox can be used to select main mixtures based on exposure and hazard data using food consumption patterns, concentration data and RPFs. The mixture selection is based on the statistical method called Sparse Nonnegative Matrix Underapproximation (SNMU) (Gillis 2013, Crepet 2019). The approach identifies the most common mixtures of substances within an assessment group to which a particular population is exposed. The results can be used for prioritisation and refinement of the mixture risk assessment. When inspection of group membership, toxicity- or exposure-related data reveals uncertainties for the substances identified in the most common mixture, more data can be collected for these substances. Refinement could include better toxicological data for assigning assessment group memberships and for RPFs and/or refinement of the exposure assessment. Since the risk is modelled based on dose addition, a refinement could also be to test the specific mixture that has been identified (i.e. not the individual substances) and determine whether application of the dose addition model overestimates the risk.

EuroMix toolbox

The module Exposure mixtures in the EuroMix toolbox is used for selection of mixtures and is based on the SNMU method.

8 Tiering approaches

Tiering in mixture risk assessment refers to the process in which different steps of the assessment can be performed using simple, conservative approaches at lower tiers and more advanced approaches requiring more data at higher tiers (EFSA 2019a, OECD 2018a). If a conservative lower tier assessment indicates that the MOE is sufficiently protective, the assessment does not have to be refined and proceed to a higher tier. Tiering approaches apply to all the different steps in the mixture risk assessment, including grouping of substances into assessment groups, hazard assessment, exposure assessment and risk characterisation. Different tiers can be used at different steps in the same assessment, e.g. a low tier approach for hazard and a high tier approach for exposure, dependent on the need for refinement and the data availability. The EuroMix methodology and tools provide possibilities to perform the assessment at different tiers (see Annex A9).

Annexes

Annex A8 describes possibilities for different tiers using the EuroMix methodology and tools.

9 Uncertainty analysis

Many of the uncertainties in a mixture risk assessment are comparable to those in risk assessment of single substances (EFSA 2018a). Mixture-specific uncertainty analysis is described in the EFSA guidance (EFSA 2019a). Uncertainties are related to e.g. the grouping of substances into assessment

groups, estimation of RPFs, missing toxicity data for included substances, missing exposure data for included substances, censored data below detection limits for concentration data of substances, choice of dose addition model and potential interactions (synergism/antagonism).

The uncertainty analysis should identify and describe the uncertainties in the different steps in a mixture risk assessment. The identified uncertainties should be quantified if possible. Annex A10 provides a template that can be used for listing the identified uncertainties and describing how they will be analysed, qualitatively or quantitatively. In the EuroMix toolbox, uncertainties related to data or other types of input for the assessment can be modelled in probabilistic 2D Monte Carlo simulations. Uncertainties can be modelled in several steps in the assessment and can be propagated to the final risk characterisation. The overall uncertainty of the risk can be visualised in plots. Annex A11 lists the possibilities of uncertainty modelling in the EuroMix toolbox.

Annexes

Annex A9 provides a template for the uncertainty analysis.
Annex A10 lists the possibilities for uncertainty analysis in the EuroMix toolbox.
Annex B8 provides an example of a filled-in template for uncertainty analysis.

10 References

AOP wiki Available from: https://aopwiki.org

ATSDR 2018. Framework for assessing health impacts of multiple chemicals and other stressors (Update). US agency for toxic substances and disease registry. Division of toxicology. Available from: http://www.atsdr.cdc.gov/interactionprofiles/ip-ga/ipga.pdf

Browse Available from: https://secure.fera.defra.gov.uk/browse

Beronius A, Molander L, Zilliacus J, Rudén C, Hanberg A. 2018. Testing and refining the Science in Risk Assessment and Policy (SciRAP) web-based platform for evaluating the reliability and relevance of *in vivo* toxicity studies. J Appl Toxicol. 38:1460-1470

Cramer GM, Ford RA, Hall RL. 1978. Estimation of toxic hazard--a decision tree approach. Food Cosmet Toxicol. 16:255-76

Crépet A, Vanacker M, Sprong C, de Boer W, Blaznik U, Kennedy M, Anagnostopoulos C, Christodoulou DL, Ruprich J, Rehurkova I, Domingo JL, Hamborg Jensen B, Metruccio F, Moretto A, Jacxsens L, Spanoghe P, Senaeve D, van der Voet H, van Klaveren J. 2019. Selecting mixtures on the basis of dietary exposure and hazard data: application to pesticide exposure in the European population in relation to steatosis. Int J Hyg Environ Health. 222:291-306

Dudzina T, Delmaar C, Biesetrbos J, Bakker M, Bokkers B, Scheeoers P, van Engelen J, Hungerbuehler K, von Goetz N. 2015. The probabilistic aggregate consumer exposure model (PACEM): validation and comparison to a lower-tier assessment for the cyclic siloxane D5. Environ Int. 79:8-16.

EFSA 2010. Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal 8(6):1637

EFSA 2011. Submission of scientific peer-reviewed open literature for the approval of pesticide active substances under Regulation (EC) No 1107/2009. EFSA Journal 9(2):2092

EFSA 2012. Guidance on the Use of Probabilistic Methodology for Modelling Dietary Exposure to Pesticide Residues. EFSA Journal 10(10):2839

EFSA 2014. Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment. EFSA Journal 12(6):3734

EFSA 2017a. Scientific Opinion on the guidance on the use of the weight of evidence approach in scientific assessments. EFSA Journal 15(8):4971

EFSA 2017b. EFSA guidance - Update: use of the benchmark dose approach in risk assessment. EFSA Journal 15(1):4658

EFSA 2018a. Guidance on Uncertainty Analysis in Scientific Assessments. EFSA Journal 16(1):5123

EFSA 2018b. Draft EFSA Scientific report Establishment of cumulative assessment groups of pesticides for their effects on the nervous system. Available from: http://www.efsa.europa.eu/en/consultations/call/180508-0

EFSA 2018c. Guidance on the use of the Threshold of Toxicological Concern approach in food safety assessment. Draft 2018. Available from: http://www.efsa.europa.eu/sites/default/files/consultation/consultation/181112-d.pdf

EFSA 2018d. The 2016 European Union report on pesticide residues in food. EFSA Journal 16(7):5348

EFSA 2019a. Guidance on harmonised methodologies for human health, animal health and ecological risk assessment of combined exposure to multiple chemicals. EFSA Journal 17(3):5634

EFSA 2019b. Draft EFSA Scientific report Establishment of cumulative assessment groups of pesticides for their effects on the thyroid. Available from: https://www.efsa.europa.eu/sites/default/files/consultation/consultation/190213-ag.pdf

EFSA Glossary. Available from: http://www.efsa.europa.eu/en/glossary-taxonomy-terms

EFSA WHO 2016. Review of the Threshold of Toxicological Concern (TTC) approach and development of new TTC decision tree. EFSA supporting publication 2016: EN-1006 Available from: https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/sp.efsa.2016.EN-1006

EuroMix 2016a. Deliverable 2.1. Kyriakopoulou K, Nikolopoulou D, Machera K, Peijnenburg A, Louca-Christodoulou D, Hadjiloizou P, Beronius A, Håkansson H, Hanberg A, Eberini I, Rorije E. Report describing cumulative assessment groups for a broad range of chemicals, based on information extracted from (literature) databases. Available from: DOI 10.5281/zenodo.1065326

EuroMix 2016b. Deliverable 2.2. Cotterill J, Eberini I, Rorije E, Peijnenburg A, Price N, Ridgway C. Report on the use of in-silico methods for the prioritisation of substances and mixtures thereof. Available from: https://cordis.europa.eu/project/rcn/193181/results/en

EuroMix 2017a. Deliverable 2.3. Rorije R, Eberini I, Peijnenburg A. Report on the use of the Threshold of Toxicological Concern (TTC) for the prioritisation of substances and mixtures thereof. Available from: https://cordis.europa.eu/project/rcn/193181/results/en

EuroMix 2017b. Deliverable 3.1. Luckert C, Braeuning A, Lampen A, Bovee T, van der Ven L. CAG-specific PCR arrays. Available from: https://cordis.europa.eu/project/rcn/193181/results/en

EuroMix 2018a. Deliverable 3.2. Luckert C, Braeuning A, de Sousa G, Durinck S, Katsanou ES, Konstantinidou P, Machera K, Milani ES, Peijnenburg A, Rahmani R, Rajkovic A, Rijkers D, Spyropoulou A, Stamou M, Stoopen G, Sturla SJ, Wollscheid B, Zucchini-Pascal N, Lampen A. Adverse Outcome Pathway-Driven Analysis of Liver Steatosis in Vitro: A Case Study with Cyproconazole. Available from: https://cordis.europa.eu/project/rcn/193181/results/en

EuroMix 2018b. Deliverable 4.2. Bois FY, Tebbya C, Brochota C. EuroMix PBPK model for combined exposures. Available from: https://cordis.europa.eu/project/rcn/193181/results/en

Garcia-Hidalgo E, Schneider D, von Goetz N, Delmaar C, Siegrist M, Hungerbühler K. 2018. Aggregate consumer exposure to isothiazolinones via household care and personal care products: Probabilistic modelling and benzisothiazolinone risk assessment, Environment International 118:245-256.

Gillis N, Plemmons RJ. 2013. Sparse nonnegative matrix underapproximation and its application to hyperspectral image analysis. Lin. Algebra Appl. 438:3991–4007.

IPCS 2009. Environmental Health Criteria 240. Principles and Methods for the Risk Assessment of Chemicals in Food. Chapter 5, p21. Available from: https://www.who.int/foodsafety/publications/chemical-food/en/

Karrer C, de Boer W, Delmaar C, Cai Y, Crépet A, Hungerbühler K, von Goetz N. 2019. Linking probabilistic exposure and pharmacokinetic modeling to assess the cumulative risk from the bisphenols BPA, BPS, BPF, and BPAF for Europeans. Submitted

Kennedy MC and Butler Ellis MC. 2016. Probabilistic modelling for bystander and resident exposure to pesticides using the Browse software. Biosystems Engineering 154:105-121.

Kennedy MC, Garthwaite DG, de Boer WJ and Kruisselbrink JW. 2019. Modelling aggregate exposure to pesticides from dietary and crop spray sources in UK residents. Environmental Science and Pollution Research. Available from: https://doi.org/10.1007/s11356-019-04440-7

Luckert C, Braeuning A, de Sousa G, Durinck S, Katsanou ES, Konstantinidou P, Machera K, Milani ES, Peijnenburg A, Rahmani R, Rajkovic A, Rijkers D, Spyropoulou A, Stamou M, Stoopen G, Sturla SJ, Wollscheid B, Zucchini-Pascal N, Lampen A. 2018. Adverse Outcome Pathway-Driven Analysis of Liver Steatosis in Vitro: A Case Study with Cyproconazole. Chem. Res. Toxicol. 31, 784−798

MCRA 9 2019. Available from: https://mcra-test.rivm.nl/EuroMix/WebApp/manual/index.html

Meek ME, Boobis AR, Crofton KM, Heinemeyer G, Raaij MV, Vickers C. 2011. Risk assessment of combined exposure to multiple chemicals: A WHO/IPCS framework. Regul Toxicol Pharmacol. 60:S1-S14.

Munro IC, Ford RA, Kennepohl E, Sprenger JG. 1996. Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. Food Chem Toxicol. 34:829-67.

NTP 2015. OHAT Risk of bias rating tool for human and animal studies. Available from: https://ntp.niehs.nih.gov/pubhealth/hat/review/index-2.html

OECD 2016. Guidance document for the use of adverse outcome pathways in developing integrated approaches to testing and assessment (IATA). ENV/JM/MONO(2016)67 Available from: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2016)67/&doclanguage=en

OECD 2018a. Considerations for assessing the risks of combined exposure to multiple chemicals, Series on testing and assessment No. 296, Environment, Health and Safety Division, Environment Directorate Available from: http://www.oecd.org/chemicalsafety/risk-assessment/considerations-for-assessing-the-risks-of-combined-exposure-to-multiple-chemicals.pdf

OECD 2018b. OECD Users' handbook supplement to the guidance document for developing and assessing AOPs Available from: https://www.oecd-ilibrary.org/environment/users-handbook-supplement-to-the-guidance-document-for-developing-and-assessing-adverse-outcome-pathways_5jlv1m9d1g32-en

Proast. Available from: https://www.rivm.nl/en/proast

RIVM, ICPS, ANSES, 2013.Toxicological data analysis to support grouping of pesticide active substances for cumulative risk assessment of effects on liver, on the nervous system and on

reproduction and development. EFSA Supporting Publications. 2013:EN-392. Available from: https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/sp.efsa.2013.EN-392

RIVM, ICPS, ANSES, 2016. Toxicological data collection and analysis to support grouping of pesticide active substances for cumulative risk assessment of effects on the nervous system, liver, adrenal, eye, reproduction and development and thyroid system. EFSA supporting publication 2016:EN-999. 184 pp. Available from: https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/sp.efsa.2016.EN-999

SciRAP. Available from: http://www.scirap.org

Slob W 2002 Dose-response modeling of continuous endpoints. Toxicol Sci 66:298-312

Solomon KR, Wilks MF, Bachman A, Boobis A, Moretto A, Pastoor TP, Phillips R, Embry MR. 2016. Problem formulation for risk assessment of combined exposures to chemicals and other stressors in humans. Crit Rev Toxicol 46:835-844

US EPA 2012. Benchmark dose technical guidance. EPA/100/R-12/001 Available from: https://www.epa.gov/sites/production/files/2015-01/documents/benchmark_dose_guidance.pdf

Vanacker M, Quindroit P, Mandin C, Glorennec P, Brochot C, Crépet A. 2019. Cumulative risk assessment to pyrethroids for neurotoxic effect through aggregate exposure and comparison of metabolite concentrations for the French adult population. To be submitted

van der Voet H and Slob W. 2007. Integration of probabilistic exposure assessment and probabilistic hazard characterization. Risk Analysis, 27: 351-371

van der Voet H, van der Heijden GWAM, Bos PMJ, Bosgra S, Boon, PE, Muri SD and Brüschweiler BJ. 2009. A model for probabilistic health impact assessment of exposure to food chemicals. Food and Chemical Toxicology, 47: 2926-2940

WHO 2018. Guidance document on evaluating and expressing uncertainty in hazard characterization– 2nd edition. Geneva: World Health Organization; 2018 Available from: https://apps.who.int/iris/handle/10665/259858

11 Glossary

Acceptable daily intake	The amount of a substance in food or drinking water that can be consumed over a lifetime without presenting an appreciable risk to health (EFSA Glossary).
Aggregate exposure	Exposure to the same substance from multiple sources and by multiple routes (EFSA 2019a).
Antagonism	Toxicological interaction in which the combined biological effect of two or more substances is less than expected on the basis of dose addition or response addition (EFSA 2019a).
Assessment group	Chemical substances that are treated as a group by applying a common risk assessment principle (e.g. dose addition) because these components have some characteristics in common (i.e. the grouping criteria) (EFSA 2019a).
Combined exposure	Exposure to multiple substances by a single route and exposure to multiple substances by multiple routes, from one or multiple sources of release and/or use(s) (OECD 2018a).
Combined hazard	Hazard from multiple substances by a single route or from multiple substances by multiple routes, from one or multiple sources of release and/or use(s) (OECD 2018a).
Component-based approach	An approach in which the risk of combined exposure to multiple substances is assessed based on exposure and effect data of the individual substances (EFSA 2019a).
Dose addition	A component-based model in which the components are treated as if having a similar action. The components may vary in toxic potency. Components contribute to the combined effect relative to the ratio between their concentration and toxic potency (EFSA 2019a).
Health-based guidance value	A numerical value derived by dividing a point of departure by a composite uncertainty factor to determine a level that can be ingested over a defined time period (e.g. lifetime or 24 h) without appreciable health risk (EFSA 2019a).
Index substance	The substance used as the point of departure for standardising the common toxicity of the substances of an assessment group (EFSA 2019a).
Interaction	Combined effects that differ from an explicit null model, i.e. dose and/or response addition. Interactions are categorised as less than additive (antagonism, inhibition, masking) or greater than additive (synergism, potentiation) (EFSA 2019a).
Line of evidence	A set of evidence of similar type (EFSA 2017a).
Margin of Exposure	Ratio of (a) a reference point of toxicity to (b) the estimated exposure dose (EFSA 2019a).
Mixture	Any combination of two or more substances that may contribute to effects regardless of source and spatial or temporal proximity (EFSA 2019a).
Mixture risk assessment	In this handbook synonymous to Risk assessment of combined exposures.
Point of Departure	The point on a dose–response curve established from experimental data used to derive a safe level. Point of Departures include LOAEL/NOAEL or benchmark dose lower confidence limit (BDML), used

	to derive a reference value or Margin of Exposure in human health risk assessment (EFSA glossary, EFSA 2019a).
Relative potency factor	Approach uses toxicity data for an index substance in an assessment group to determine potency-adjusted exposure data for substances in the mixture assuming similarity of mode of action between individual substances in the mixture (EFSA 2019a).
Response addition	A component-based model in which the components are treated as if having independent or dissimilar action, i.e. by following the statistical concept of independent random events (EFSA 2019a).
Risk assessment of combined exposures	Risk assessment of exposure to multiple substances by a single route and risk assessment of exposure to multiple substances by multiple routes, from one or multiple sources of release and/or use(s) (OECD 2018a).
Synergism	Toxicological interaction in which the combined biological effect of two or more substances is greater than expected on the basis of dose addition or response addition (EFSA 2019a).
Tolerable daily intake	Estimate of the amount of a substance in food or drinking water which is not added deliberately (e.g. contaminants) and which can be consumed over a lifetime without presenting an appreciable risk to health (EFSA glossary).
Toxicodynamics	Process of interactions of toxicologically active substances with target sites in living systems, and the biochemical and physiological consequences leading to adverse effects (EFSA 2019a).
Toxicokinetics	Process of the uptake of substances, by the body, the biotransformations they undergo, the distribution of the parent substances and/ or metabolites in the tissues, and their elimination from the body over time (EFSA 2019a).
Whole mixture approach	A risk assessment approach in which the mixture is treated as a single entity, similar to single substances, and so requires dose-response information for the mixture of concern or a (sufficiently) similar mixture (EFSA 2019a).

12 Annexes A - Detailed methodology and templates

12.1 Annex A1 - Template for analysis plan

The template can be used to summarise the outcome of the problem formulation in an analysis plan. The assessor should briefly describe each item in the column “Recorded information”. The problem formulation should be viewed as an iterative process and refinements of the analysis plan may become necessary as the risk assessment progresses and information is gathered.

Problem formulation element	Description	Recorded information
Risk assessment question	Specific question to be addressed	
Description of mixture	Evidence for common toxicological effect of the mixture components	
	Evidence for co-exposure	
Conceptual model	Regulatory framework or remit	
	Substance categories	
	Exposure source(s) (e.g. food, drinking water, cosmetics, consumer products, air, soil)	
	Exposure route(s) (e.g. oral, dermal, inhalation)	
	Population group (e.g. general population, workers, school children, pregnant women, country)	
	Population age (e.g. infant, toddler, child, teen, adult, elderly)	
	Toxicological effect	
	Level of grouping (common target organ, common effect/adverse outcome or common specific mode of action/AOP)	
Methodology	Data availability for toxicity described in general terms, including the type of data	
	Data availability for exposure described in general terms, including the type of data	
	Use of AOP networks to support grouping of substances into assessment groups and/or identification of measurable effect/key events	
	Collection of toxicity data from literature	
	Generation of toxicity data	
	Grouping of substances based on toxicological considerations	
	Calculation of RPFs, choice of index substance and selection of POD	

	Extrapolation between *in vitro* and *in vivo* studies	
	Approach for dealing with lack of toxicity data	
	Collection of consumption data	
	Collection of concentration data	
	Non-detects, concentration measurements below the limit of detection	
	Approach for dealing with lack of concentration data	
	Conversion of food-as-eaten to food-as-measured and processing factors	
	Acute or chronic model for dietary exposure assessment	
	Model for dietary exposure assessment (deterministic or probabilistic)	
	Model for non-dietary exposure (if applicable)	
	Choice of model (dose addition or response addition)	
	Consider evidence for interactions (synergism or antagonism)	
	Risk metrics to be used (e.g. margin of exposure) and different percentiles of the exposure distribution to be used	
	Initial tiers used for the different parts of the assessment	
	Methodology to identify risk drivers in a lower tier assessment	
	Higher tiers to be used when lower tier assessment does not provide sufficient protection	
	Method(s) for uncertainty analysis (e.g. qualitatively or quantitatively)	
	Uncertainty modelling for quantitative uncertainty analysis of data and other types of input	
	Types of expertise needed	
	Number of persons needed for the assessment	
	Other resources needed	
	Estimated timeframe for the assessment	
	Stakeholder consultation and peer review	

12.2 Annex A2 - Organisation of information on AOP networks and toxicity data in the EuroMix toolbox

The format for toxicity data entered into the EuroMix toolbox is based on the concept of KEs and KERs in an AOP network. This is how all relevant data for specific responses are collected, described and linked to each other and to specific effects (KEs). Unique IDs are required to identify all elements. The data is entered in specific tables, i.e. Excel sheets. The tables contain columns both for required data and for optional data that can be used for reference. Detailed information is available in the EuroMix toolbox manual (MCRA 9 2019).

The AOP networks table describes the AOP network. Information on the adverse outcome of the AOP network is required and the table can also include reference to corresponding AOPs described in the AOP wiki.

The Effects table describes the effects that are focus for the mixture risk assessment. The table only requires the ID of the effect, but it is recommended that the effect is described using the harmonised format of KE descriptions in the AOP wiki (https://aopwiki.org/) to the extent possible (OECD 2018). As such, Effect descriptions should include descriptors for the Process, Object, Action, Cell and Organ relevant for the KE. If the KE has been previously described in the AOP wiki, descriptors can simply be copied from the KE description. If the effect does not correspond to a previously described KE in the wiki, attempts should be made to follow the wiki format as closely as possible. If the effect has been described as a KE in the AOP wiki, it is recommended to include the KE number assigned in the wiki.

The Effect relations table describes the relationships between the different up-stream and down-stream effects that correspond to KERs in the AOP network. Each Effect Relation should be described using the ID for the AOP network, as well as the IDs for the up-stream and down-stream effect. In other words, Effect Relations are not assigned their own specific IDs. If the Effect Relation has been described as a specific KER in the AOP wiki, it is recommended to include the KER number assigned in the wiki. If the Effect Relation has not been described in the AOP wiki but is supported by other data, references to relevant publications should be included.

The Test systems table describes the *in vivo* or *in vitro* system where responses are measured. The table only requires the ID of the test system, but it is recommended that also the type of test system (i.e. cell line, primary cells, organ culture, *in vivo*) as well as species, strain and organ are described. The route of exposure can also be described.

The Responses table describes the specific responses measured in the test system. The table requires information on the test system used and the response type, i.e. is the result expressed as continuous or quantal data. Reference to any guideline for the assay can also be included.

The Effect representations table describes how the responses are also linked to specific effects. The table requires information on the effect and the corresponding response. Benchmark responses used for the dose response modelling should also be specified here. The benchmark response specifies the minimum response required to conclude that a substance causes an effect.

The Dose response data should be provided for each separate experiment and substance in a specific format. The Excel file should contain a worksheet with details of the experiment (named DoseResponseExperiments). The required information is the ID of the experiment, the substances and responses tested, and the unit of the dose of the substance. The dose response data is provided in separate worksheets. The worksheets must include information on the doses tested and the

results expressed either as individual data or summary data including the standard deviation or coefficient of variation and the number of measurements.

12.3 Annex A3 - Template for extraction of toxicity data for mixture risk assessment

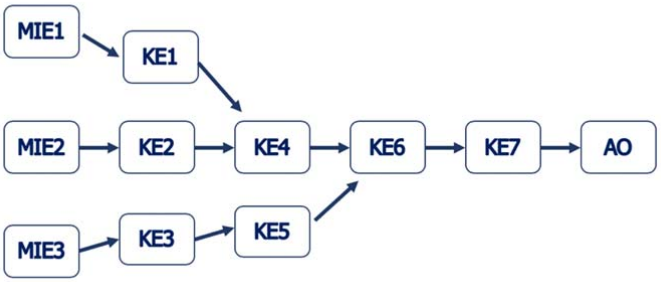
The table provides a template for extraction of toxicity data from *in vivo* and *in vitro* studies for the purpose of either grouping or relative potency factors. The table follows the data formats used in the EuroMix toolbox. The template is also available as an Excel file including drop-down lists that facilitate data extraction.

StudyID	ID number of study if used
Source	Source of study, e.g. open literature, EFSA DAR
Reference	Reference to scientific paper, or study identifier
Year	Year of study, e.g. year when study was reported or published
idSubstance	EuroMix code for substances
SubstanceName	Name of substance
SubstancePurity	Purity of the test substance, if described
GuidelineMethod	Guideline used for the study design, if given (e.g. OECD 407)
GLP-GIVIMP	Was the study performed according to GLP (good laboratory practice) or GIVIMP (good *in vitro* methods practices)
TestSystem	*In vivo* studies: Study type and/or length (e.g. rat 28-day, mouse developmental). *In vitro* studies: e.g. name of cell line or cell type. EuroMix codes for test systems can be used
TestSystemType	Type of test system
Species	Animal species tested (e.g. rat, mouse, dog)
Strain	Name of animal strain. Only relevant for *in vivo* studies.
RouteExposure	Route of exposure
TypeAdministration	Indicate the mode of administration (e.g. capsule, dietary, gavage, drinking water). Only relevant for *in vivo* studies.
TestedDoses	Listing of all the individual doses (e.g. 1, 3, 5, 10) or range of the doses used (e.g. 1-10), when individual doses are not available
DoseUnit	The dose unit of the doses (e.g. mg/kg bw, mM)
ExposureDuration	Duration of exposure to the substance in e.g. hours, days, weeks
TimePointObservation	Time point when the response was observed or measured, in cases it was not directly at the end of exposure
Effect	Effect (i.e. key events), e.g. activation of LXR receptor, liver steatosis. EuroMix codes for effects can be used
Response	Response (i.e. endpoint) measured in the study. E.g. LXR activation of reporter gene in HepG2 cells, fatty acid accumulation in liver cells. EuroMix codes for responses can be used
Sex	Sex of animals in which the response is measured/observed.
Age	The age of animals in which the response is measured/observed. E.g. Fetus, Juvenile, Adult
Generation	Generation of animals in which the response is measured/observed. E.g. F0, F1, F2
LowestResponseDose	The lowest dose for which the response can be observed. Dose unit is reported in column DoseUnit. In case response cannot be observed at any dose, state no response.
PointOfDeparture	Point of departure of the measured response in the study. Dose unit is reported in column DoseUnit. In case response cannot be observed at any dose, state no response.
PODType	The type of the point of departure, e.g. NOAEL, LOAEL, BMD
BenchmarkResponse	The benchmark response (BMR) used in case Point of departure is expressed as BMD. E.g. 5%
Relevance	Assessment of the relevance of the measured response for the specific purpose, either grouping of substances into assessment groups or calculation of relative potency factors. Scores or categories can be used to express relevance

Reliability	Assessment of the reliability of the result for grouping of substances into assessment groups or for the purpose of calculation of relative potency factors. Scores or categories can be used to express reliability
Comments	Any comments on the data

12.4 Annex A4 - Template for describing a tiered testing strategy for mixture risk assessment based on AOP networks

The template can be used to describe a tiered testing strategy for mixture risk assessment based on AOP networks. The AOP network is described with a figure and the selected *in silico* and *in vitro* assays are described in the table.



KE number in AOP network	KE name	*In silico* model/*in vitro* assay for measuring the KE	Relevance of the *in silico* model/*in vitro* assay	Reliability of the *in silico* model/*in vitro* assay	Availability and feasibility of *in silico* model/*in vitro* assay	Information provided by the *in silico* model/*in vitro* assay for the mixture risk assessment (e.g. for grouping, RPFs and/or prioritisation for further testing)
MIE1						
MIE2						
KE1						
KE2						
KE3						
KE4						
KE5						
KE6						
KE7						
AO						

12.5 Annex A5 - Template for organising data for grouping

The template describes how to organise evidence to be used for grouping into assessment groups. The evidence is organised according to different lines of evidence, i.e. each key event and *in silico*, *in vitro*, *in vivo* or human studies. Rows are added in the table for each piece of evidence.

Substance	**Key event in the AOP network (organised according to MIE, intermediate KEs, AO)**	**Study type (organised according to *in silico*, *in vitro*, *in vivo* data, human study)**	**Assay (specific assay used)**	**Main study result (e.g. positive, negative, BMDL, NOAEL)**	**Reliability (low, medium, high)**	**Relevance (low, medium, high)**
	MIE	*In silico*				
		In vitro				
		In vivo				
		Human				
	Each intermediate KE	*In silico*				
		In vitro				
		In vivo				
		Human				
	AO	*In silico*				
		In vitro				
		In vivo				
		Human				

12.6 Annex A6 - Methodology to derive and use probabilities of assessment group membership

Substances are grouped into assessment groups based on information on toxicity. The available information can be *in silico* data from QSAR or molecular docking, *in vitro* data from studies in cells, *in vivo* data from animal studies or human data from epidemiological studies. All data should be considered, and assessed for reliability and relevance. Each substance is given a probability between 0-1 for belonging to the assessment group. 0 indicates that it is certain that the substance does not belong to the group, 1 indicates that it is certain that the substance does belong to the group and values between 0-1 indicates the uncertainty about the membership. Methods for deriving probabilities can be expert knowledge elicitation based on all available data or information from *in silico* data when no *in vitro*, *in vivo* or human data is available.

The EuroMix toolbox can consider the probability of the assessment group membership. This is done in the "Active substances module". Each substance will have an assessment membership probability (0-1), derived from expert knowledge elicitation or *in silico* data. Depending on the required model tier and available data, these probabilities can be specified directly using input data or calculated in the toolbox. Alternative methods are described in the user manual. These include options to set a prior probability of membership (e.g. expert assessment independent of the data) and to account for the specificity and sensitivity of individual QSAR models within the calculation. The assessment membership probability is used in probabilistic modelling using 2D Monte Carlo simulation. The modelling is based on a large number of iterations of the Monte Carlo simulation. Substances with a probability of 0 will not be included in any of the simulations, whereas substances with a probability of 1 will be included in every iteration. Substances with a probability between 0 and 1 will be included in each iteration with the probability of the group membership.

12.7 Annex A7 - Methodology to design and analyse mixture experiments

The following methodology can be used to design and analyse mixture experiments, using binary mixtures.

- The mixture experiment is performed using doses with equal potency. Therefore relative potency information for the individual substances is required to design the mixture experiment. It is preferable that the individual substances first are tested in the selected assay and the relative potency is calculated. In case this is not possible relative potency derived from other assays can be used, e.g. relative potency from *in vitro* assays can be used to design *in vivo* testing. The relative potency of the individual substances can be identified using the benchmark dose method (see section 5.5.1).
- The mixture experiment should include several doses of each individual substance and of the mixture of the two substances. The mixture doses should be equipotent to the individual doses (see Table 1).
- The result is analysed using the benchmark dose method in a combined analysis of the individual substances and the mixture. The fitted dose-response curve is analysed. If dose addition applies to the mixture, its data points will not show a systematic deviation of the curve derived from the single compounds. In cases of less or more than dose addition, the mixture data points will show a shift to either the right or the left, respectively, of the single compounds curve (see Fig. 1).

Table 1. Example of dose selection for mixture experiments. The selected doses for the single substances and the mixture are equipotent, taking into account the RPFs.

Substance 1 RPF=1	Substance 2 RPF=5
Dose in e.g. μM	
0	
1	
2	
4	
8	
16	
32	
	0
	1/5=0.2
	2/5=0.4
	4/5=0.8
	8/5=1.6
	16/5=3.2
	32/5=6.4
0	0
1/2=0.5	1/5/2=0.1
2/2=1	2/5/2=0.2
4/2=2	4/5/2=0.4
8/2=4	8/5/2=0.8
16/2=8	16/5/2=1.6
32/2=16	32/5/2=3.2

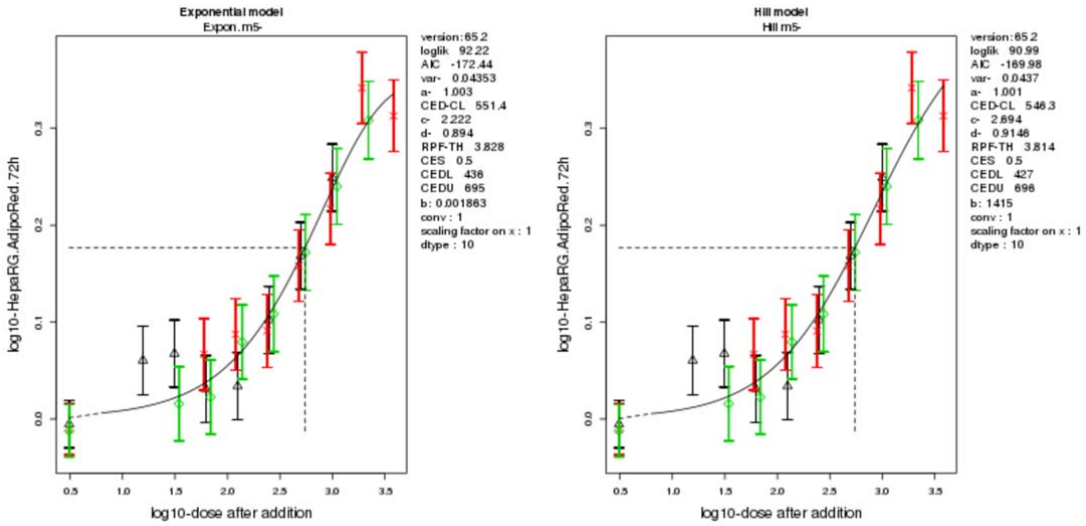


Figure 1. Analysis of mixture experiments using benchmark dose method. The two graphs shows two fitted models using Proast software. The black triangles and red crosses show the data points for the single substances and the green diamonds for the mixture.

12.8 Annex A8 - Possibilities for different tiers using EuroMix methodology and tools

The table describes possibilities for different tiers using EuroMix methodology and tools for hazard and exposure assessment.

Step in assessment	**Possibilities for tiering**
Hazard assessment	
Grouping into assessment groups	At lower tier all substances that have a common target organ can be grouped forming a large assessment group and at higher tiers substances can be grouped based on a common effects/adverse outcome (see section 5.4)
Potency of substances	At lower tiers can the lowest POD for the substances in the group be used for all substances and at higher tiers can substance-specific PODs used (see section 5.5)
Effect used for RPFs	At lower tiers can the critical effect (lowest POD for substance for any effect) be used and at higher tiers can the specific effect that is the basis for grouping be used (see section 5.5)
Missing toxicity data	At lower tiers can Munro NOAELs be used and at higher tiers can new toxicity data generated (see section 5.5.6)
Exposure assessment	
Consumption data	At lower tiers can physiological limits of consumption be used and at higher tiers individual data from consumption surveys (see section 6.1)
Concentration data	At lower tiers can maximum permitted levels be used and at higher tiers data from representative monitoring studies (see section 6.1)
Missing concentration data	At lower tiers can maximum permitted levels be used and at higher tiers can new concentration data generated (see section 6.2)

12.9 Annex A9 - Template for uncertainty analysis

The template can be used for uncertainty analysis of mixture risk assessment performed using the EuroMix methodology and tools. The assessor should first consider which uncertainties could be related to each aspect and list them in the column “Identified uncertainties”. The assessor should describe in the column “Analysis of uncertainty” how each identified uncertainty will be analysed, qualitatively or quantitatively. Some uncertainties related to data or other types of input can be modelled probabilistically using 2D Monte Carlo simulations in the EuroMix toolbox. This can be described in the column “Analysis of uncertainty “.

Aspect	**Identified uncertainties**	**Analysis of uncertainty (qualitative or quantitative)**
Grouping of substances into assessment groups		
Choice of toxicity data to derive POD		
Calculation of RPFs		
Extrapolation between *in vitro* and *in vivo* studies		
Lack of toxicity data		
Consumption data		
Concentration data		
Non-detects, concentration measurements below the limit of detection		
Lack of concentration data		
Conversion of food-as-eaten to food-as-measured and processing factors		
Other (non-dietary) routes of exposure		
Use of the dose addition model		
Possible interactions (synergism or antagonism)		

12.10 Annex A10 - Uncertainty modelling in the EuroMix toolbox

Part of the uncertainties related to data or other type of input can be modelled in the EuroMix toolbox using probabilistic 2D Monte Carlo. The table lists the possibilities for uncertainty modelling. There are three basic approaches:

1. Use additional sets of data to characterise the uncertainty distribution of input data (uncertainty sets). Typically 100-1000 sets are provided, e.g. from external modelling programs by the user. In some cases, however, sets that are already available can be used (imputation of hazard characterisations and exposures).
2. Bootstrap the primary datasets. This approach requires no further user specification.
3. Parametric approach: For several entities in the toolbox parametric distributions have been implemented to characterise uncertainty. These typically require that parameters such as probabilities, upper values or coefficients of variation are specified. In the case of concentration uncertainty in dietary exposures, parametric uncertainty based on fitted concentration models can be used.

Approach	**Module in EuroMix toolbox**	**User input needed**	**Description of values used in uncertainty iterations**
Sets of data	Points of departure	yes	Alternative POD values, user-specified data in table PointsOfDepartureUncertain
	Dose response models	no/yes	Alternative BMD values, calculated or user-specified data in table DoseResponseModelBenchmarkDosesUncertain
	Hazard characterisation	no	Imputation of missing hazard characterisations with sampled values from the Munro NOEL collections or other selected NOAELs
	Relative potency factors	no/yes	Alternative RPF values, calculated or user-specified data in table RelativePotencyFactorsUncertain
	Non-dietary exposures	yes	Alternative non-dietary exposures, user-specified data in table NonDietaryExposuresUncertain
	Dietary exposures	no	Imputation of missing exposures with sampled values from other selected dietary exposures
Bootstrap	Dose response data	no	Parametric bootstrap sample of dose response data
	Consumptions	no	Non-parametric bootstrap sample of consumption data
	Concentrations	no	Non-parametric bootstrap sample of concentration data
Parametric	Concentrations	no	Sample from the estimated uncertainty distributions of the fitted Concentration models
	QSAR memberships	yes	Calculate assessment group membership probabilities of substances using sensitivity and specificity of QSAR models (user-specified data in table QSARMembershipModels)
	Active substances	no/yes	Sample from Bernoulli distribution (0 or 1) for assessment group membership of substances using membership probabilities (calculated or user-specified data in table AssessmentGroupMemberships)
	Inter-species factors	yes	Sample from lognormal distribution using geometric mean and geometric standard deviation (user-specified data in table InterSpeciesModelParameters)
	Intra-species factors	yes	Sample from chi-square distribution to define the width of the lognormal variability-between-individuals distribution, using lower and upper values for the sensitivity factor for a sensitive human individual (user-specified data in table IntraSpeciesModelParameters)
	Kinetic models	yes	Sample from log-normal or log-logistic distributions for parameter values of the kinetic model, using coefficient of variation (user-specified data in table KineticModelInstanceParameters)

	Processing factors	yes	Sample from log-normal, log-logistic or chi-square distributions, using upper values of the median processing factor and/or upper values of the upper limit for a variable processing factor (user-specified data in table ProcessingFactors)

13 Annexes B - Examples

13.1 Annex B1- Example of an analysis plan for mixture risk assessment

The example of an analysis plan for mixture risk assessment is based on work performed in the EuroMix project during development of the EuroMix methodology and tools, but should not be seen a real mixture risk assessment.

Problem formulation element	Description	Recorded information
Risk assessment question	Specific questions to be addressed	What is the risk for the adult population in the Netherlands to develop liver steatosis due to combined dietary exposure to pesticide residues in food?
Description of mixture	Evidence for common toxicological effect of the mixture components	Liver steatosis is identified as an effect in *in vivo* studies of several pesticides (RIVM, ICPS, ANSES, 2013, 2016)
	Evidence for co-exposure	Monitoring studies show that European food contains low levels of several pesticide residues (EFSA 2018d)
Conceptual model	Regulatory framework or remit	Not applicable
	Substances categories	Pesticide residues
	Exposure source(s) (e.g. food, drinking water, cosmetics, consumer products, air, soil)	Food and drinking water
	Exposure route(s) (e.g. oral, dermal, inhalation - specify as needed)	Oral
	Population group (e.g. general population, workers, school children, pregnant women, country)	General population in the Netherlands
	Population age (infant, toddler, child, teen, adult, elderly)	Adults
	Toxicological effect	Liver steatosis
	Level of grouping (common target organ, common effect/adverse outcome or common specific mode of action/AOP)	Common effect/adverse outcome
Methodology	Data availability for toxicity described in general terms, including the type of data	Regulatory in vivo toxicity studies reported in EFSA draft assessment reports, other assessment reports from international bodies. In vivo studies in scientific papers. Study characteristics, NOAEL, LOAEL, but in most cases not dose-response data. Results from QSAR modelling and in vitro studies.

	Data availability for exposure described in general terms, including the type of data	Consumption data from national dietary surveys in Netherlands. Concentration data from European monitoring studies.
	Use of AOP networks to support grouping of substances into assessment groups and/or identification of measurable effect/key events	AOP networks for liver steatosis used to support grouping and identify measurable key events
	Collection of toxicity data from literature	Literature search to identify toxicity data from reports and scientific papers, selection of studies on included substances and effect included to the AOP network, assessing reliability of scientific papers using SciRAP tool.
	Generation of toxicity data	In case relevant and reliable toxicity data is not identified from literature, in silico data and in vitro data for selected KEs in the AOP network is generated.
	Grouping of substances based on toxicological considerations	Grouping at the level of common effect, liver steatosis, using in silico, in vitro and in vivo data using two experts. Grouping reported as included/not included.
	Calculation of RPFs, choice of index substance and selection of POD	Benchmark dose modelling is performed when dose-response data is available. NOAELs are used in other cases. Index substance selected based on criteria in EuroMix handbook. Most conservative POD selected.
	Extrapolation between *in vitro* and *in vivo* studies	In vitro POD extrapolated to in vivo POD using kinetic factors.
	Approach for dealing with lack of toxicity data	Imputation of missing PODs using Munro TTC values.
	Collection of consumption data	Food consumption surveys for the adult population in Netherlands, using 24 h-recall on 2 non-consecutive days.
	Collection of concentration data	Concentration data from European control and monitoring programmes, using objective or selective sampling.
	Non-detects, concentration measurements below the limit of detection	Use zero value.
	Approach for dealing with lack of concentration data	Extrapolation of measured data from other foods and using MRLs.
	Conversion of food-as-eaten to food-as-measured and processing factors	Use of conversion table based on recipes from the Netherlands and published processing factors.

	Acute or chronic model for dietary exposure assessment	Chronic since liver steatosis is caused by long-term exposure
	Model for dietary exposure assessment (deterministic or probabilistic)	Probabilistic model
	Model for non-dietary exposure (if applicable)	Only dietary exposures
	Choice of model (dose addition or response addition)	Dose addition model as the default model
	Consider evidence for interactions (synergism or antagonism)	In case data indicates interactions, a suitable approach will be discussed
	Risk metrics to be used (e.g. margin of exposure) and different percentiles of the exposure distribution to be used	MOE using different percentiles up to 99.9 percentile
	Initial tiers used for the different parts of the assessment	Grouping at the level of common effect, liver steatosis. POD for specific effect liver steatosis. Imputation of missing PODs using Munro TTC values. Consumption data from dietary surveys. Concentration data from European control and monitoring programmes. MRLs used for missing concentration data. Probabilistic exposure assessment.
	Methodology to identify risk drivers in a lower tier assessment	MOE at 99.9 percentile below 100
	Higher tiers to be used when lower tier assessment does not provide sufficient protection	Generation of hazard and concentration data.
	Method(s) for uncertainty analysis (e.g. qualitatively or quantitatively)	Uncertainties listed and quantified when possible
	Uncertainty modelling for quantitative uncertainty analysis of data and other types of input	Probabilistic exposure assessment
	Types of expertise needed	Experts on exposure assessment, toxicology and risk assessment.
	Number of persons needed for the assessment	Not applicable
	Other resources needed	EuroMix toolbox, access to toxicity, concentration and consumption data.
	Estimated timeframe for the assessment	Not applicable
	Stakeholder consultation and peer review	Not applicable

13.2 Annex B2 - Example on development and assessment of an AOP network for mixture risk assessment

This example of development and assessment of an AOP network for liver steatosis is based on work performed in the EuroMix project, but should not be seen as a final, fully developed and assessed AOP network that can be used for mixture risk assessment.

The example follows the methodology described in the EuroMix handbook.

Identification of existing published AOPs

First existing published AOPs for liver steatosis were identified in the AOP wiki https://aopwiki.org and in literature.

Six potential AOPs were identified in the AOP wiki. They are all still under development and have not been reviewed and endorsed by OECD.

- AOP 34 LXR activation leading to hepatic steatosis
- AOP 36 Peroxisomal Fatty Acid Beta-Oxidation Inhibition Leading to Steatosis
- AOP 57 AhR activation leading to hepatic steatosis
- AOP 58 NR1I3 (CAR) suppression leading to hepatic steatosis
- AOP 60 NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis
- AOP 61 NFE2L2/FXR activation leading to hepatic steatosis

Two publications describing potential AOPs were identified from literature:

- Vinken M. 2015. Adverse Outcome Pathways and Drug-Induced Liver Injury Testing. Chem Res Toxicol 28:1391–1397
- Mellor et al. 2016. The identification of nuclear receptors associated with hepatic steatosis to develop and extend adverse outcome pathways. Critical Reviews in Toxicology, 46:138-152.

The identified AOPs start at different MIEs but end at the common AO liver steatosis. An AOP network including several AOPs was postulated based on the information in the six AOPs in the AOP wiki and the two publications, see Figure 1. The AOP network is also described in Luckert 2018 and EuroMix 2018a.

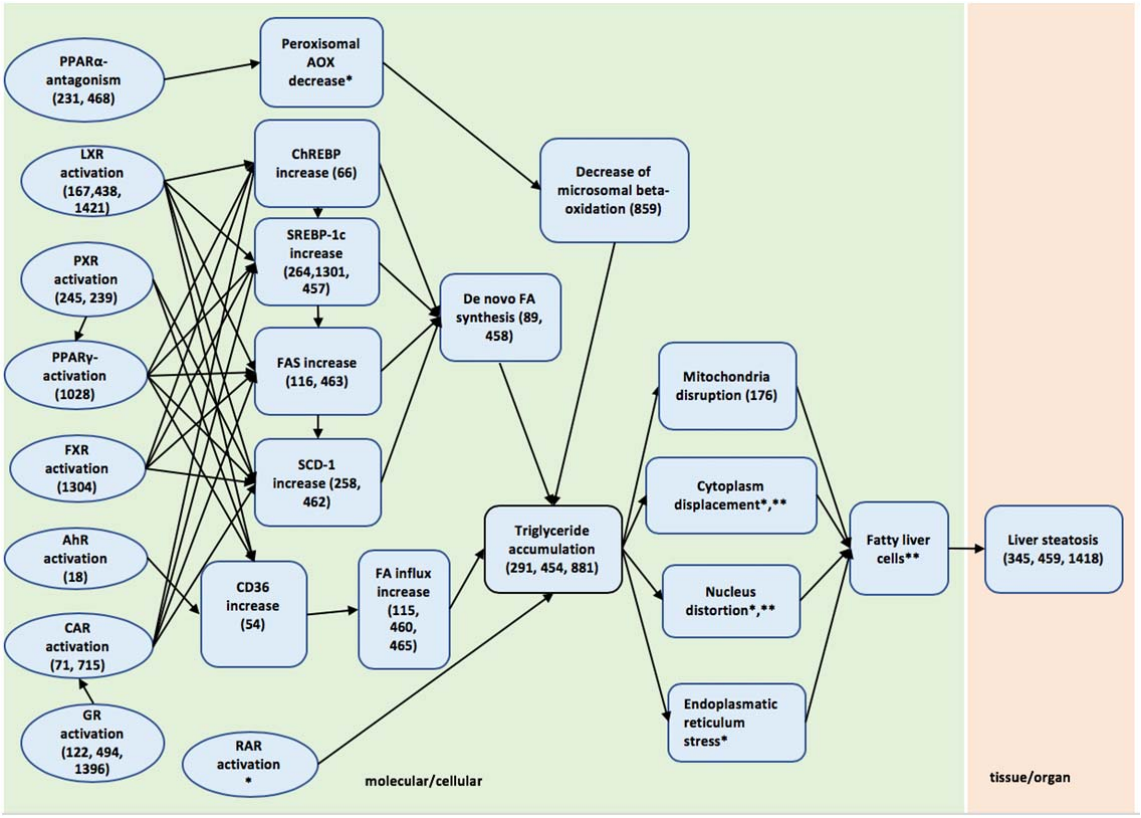


Figure 1. Postulated AOP network for liver steatosis. The ovals are MIEs and the rectangles are KEs. The rectangle "liver steatosis" is the AO. The arrows depict KERs. The numbers in the ovals/rectangles refer to KE numbers in the AOP wiki. * refers to KE not included in the AOP wiki but described in Mellor 2016. ** refers to KE not included in the AOP wiki but described in Vinken 2015.

Assessment of the postulated AOP network

Assessment of an AOP network should be performed according to the OECD AOP handbook (OECD 2018b). The assessment includes evaluation of the biological plausibility and empirical support for the KE relationships linking the KEs, as well as evidence supporting the essentiality of the KEs.
The liver steatosis AOPs in the AOP wiki have not yet been assessed and a formal assessment of the postulated AOP network for liver steatosis has not been done in the EuroMix project.

Description of the AOP network in tables for use in the EuroMix toolbox

The postulated AOP network for liver steatosis, including the MIEs, KEs, AO and KER, was described in tables (i.e. Excel sheets) for use in the EuroMix toolbox. The Excel file with the sheets are uploaded into the EuroMix toolbox. The Excel sheets contain the following information:

AOP networks table

Name	Description	Liver steatosis AOP network information
idAOPN	Unique identification code of the AOP network	AOPN-steatosis
Name	Name of the AOP network	AOPN-steatosis
Description	Additional description or label of the AOP network	Steatosis AOPnetwork
idAdverseOutcome	The identification code of the effect representing the adverse outcome of this AOP network	Steatosis-liver
RiskType	The risk type of the adverse outcome	Chronic
AOPwikiAOPs	Reference to any AOPs in the AOP wiki.	34,36,57,58,60,61
References	External reference(s) to sources containing more information about the AOP network.	EuroMix doc: "EuroMix steatosis network"

Effects table

Name	Description	Example of liver steatosis AOP network information. Similar data is included for each key event (effect)
idEffect	Unique identification code of the effect	LXR-act-liver
Name	Name of the effect	LXR-act-liver
Description	Additional description or label of the effect	Activation of Liver X receptor (LXR) signaling in liver
BiologicalOrganisation	Biological organisation of the effect: Molecular, Cellular, Tissue, Organ, Individual. This is in line with AOP wiki terminology and can be used for grouping	Molecular
KeyEventProcess	Description of AOP Key event component process. E.g., receptor signalling. This is according to the ontologies AOP wiki terminology	signaling
KeyEventObject	Description of AOP Key event component object. E.g., PPAR-alpha. This is according to the ontologies AOP wiki terminology	oxysterols receptor LXR-alpha AND oxysterols receptor LXR-beta
KeyEventAction	Description of AOP Key event component action. E.g., decreased. This is according to the ontologies AOP wiki terminology	increased
KeyEventCell	Description of AOP Key event organ. E.g., hepatocyte. This is according to the AOP wiki terminology	hepatocyte
KeyEventOrgan	Description of AOP Key event organ. E.g., liver. This is according to the AOP wiki terminology	liver
AOPwikiKE	Key event ID number in AOP wiki. Several ID	167,483,1421

	possible. Some effects might not be in the wiki, and this field will be empty	
References	External reference(s) to sources containing more information about the AOP key event.	Vinken. 2015. Chem Res Toxicol, 28:1391–1397; Mellor et al. 2016. Critical Reviews in Toxicology, 46:138-152

Effects relations table

Name	Description	Example of liver steatosis AOP network information. Similar data is included for each key event relationship (effects relation)
idAOPN	Identification code of the AOP network for which this link is defined.	AOPN-steatosis
idUpstreamKeyEvent	Identification code of the triggering effect of this relationship.	LXR-act-liver
idDownstreamKeyEvent	Identification code of the (triggered) effect of this relationship.	ChREBP-incr-liver
AOPwikiKER	Key event relationshop ID number in AOP wiki. Several ID possible. Some effect relations might not be in the wiki, and this field will be empty	174
reference	External reference(s) to sources containing more information about the effect (key event) relationships.	Vinken. 2015. Chem Res Toxicol, 28:1391–1397; Mellor et al. 2016. Critical Reviews in Toxicology, 46:138-152; EuroMix milestone 3.1

13.3 Annex B3 - Example of a tiered testing strategy for mixture risk assessment based on AOP networks

This example of development and assessment of a tiered testing strategy for liver steatosis is based on work performed in the EuroMix project, but should not be seen as a final, fully developed and assessed testing strategy that can be used for mixture risk assessment.

The example follows the methodology described in the EuroMix handbook.

The postulated AOP network for liver steatosis described in Annex B2 formed the basis for the tiered testing strategy, see figure 1.

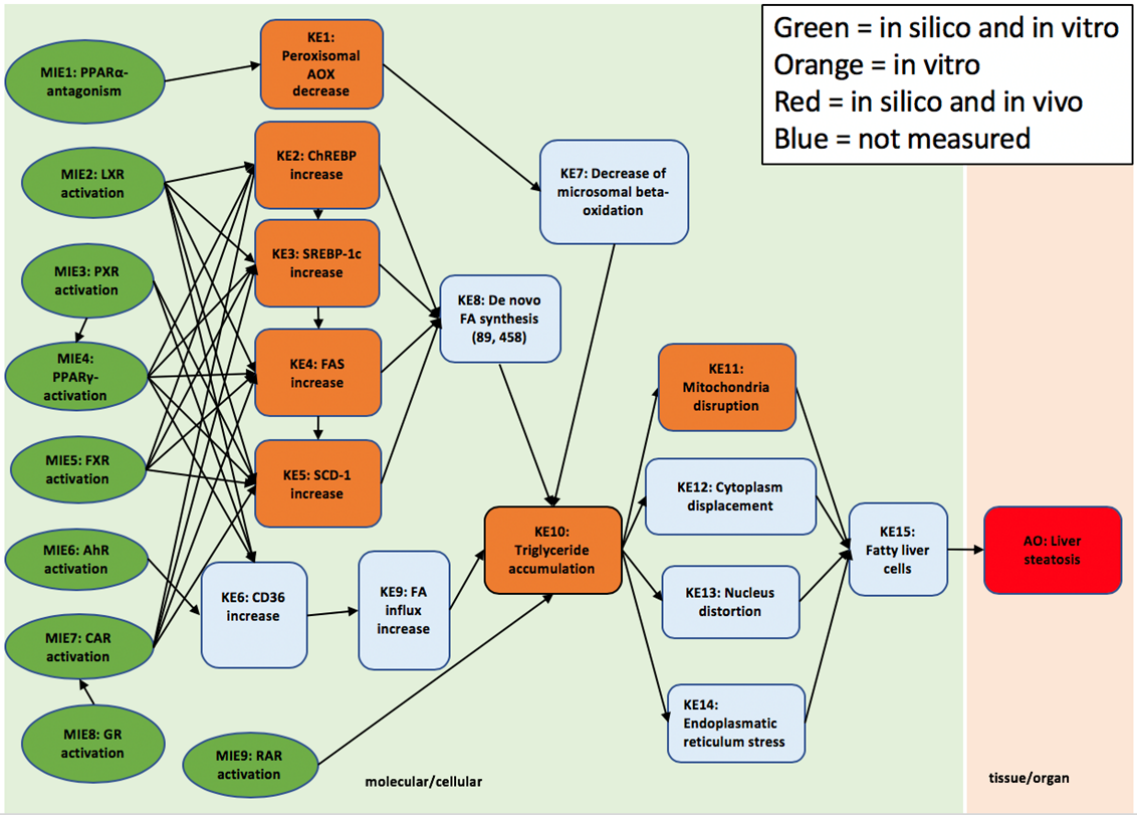


Figure 1. Tiered testing strategy for liver steatosis. For MIEs in green ovals, *in silico* models and *in vitro* assays are included. For KEs in orange rectangles, *in vitro* assays are included. For AO in red rectangle, *in silico* models and an *in vivo* test are included. For KE in blue rectangles, no assays are included.

All KEs in the postulated AOP network were considered to be potentially useful for providing information for grouping of substances into assessment groups and/or to derive RPFs. Possible *in silico* models and *in vitro* and *in vivo* assays were identified. Their relevance, reliability, availability/feasibility and usefulness were assessed, although some aspects have not yet been fully assessed. The results are summarised in the table below and also described in Luckert 2018, EuroMix 2016b, 2017b, 2018a.

KE number in AOP network	KE name	*In silico* model/*in vitro* assay for measuring the KE (EuroMix code)	Relevance of the *in silico* model/*in vitro* assay	Reliability of the *in silico* model/*in vitro* assay	Availability and feasibility of *in silico* model/*in vitro* assay	Information provided by the *in silico* model/*in vitro* assay for the mixture risk assessment (e.g. for grouping, RPFs and/or prioritisation for further testing)
MIE1	PPARalpha-antagonism-liver	Molecular docking model for hPPARalpha (MolDock-hPPARalpha)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE1	PPARalpha-antagonism-liver	PPARα activation of reporter gene in HepG2 cells (HepG2-RGA-PPARalpha)	Relevant reporter gene assay in human cell line to measure activation and antagonism of the receptor	Reliable assay based on positive and negative controls	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
MIE2	LXR-act-liver	Molecular docking model for hLXRalpha (MolDock-hLXRalpha)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE2	LXR-act-liver	LXR activation of reporter gene in HepG2 cells (HepG2-RGA-LXR)	Relevant reporter gene assay in human cell line to measure receptor activation	Reliable assay based on positive and negative controls	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
MIE2	LXR-act-liver	Molecular docking model for hLXRbeta (MolDock-hLXRbeta)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE3	PXR-act-liver	Molecular docking model for hPXR (MolDock-hPXR)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE3	PXR-act-liver	PXR activation of reporter gene in HepG2 cells (HepG2-RGA-PXR)	Relevant reporter gene assay in human cell line to measure receptor	Reliable assay based on positive and negative	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data

			activation	controls		
MIE4	PPARgamma-act-liver	OCHEM QSAR model for PPARgamma (QSAR-OCHEM-PPARgamma-Hepatotoxicity)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping
MIE4	PPARgamma-act-liver	Molecular docking model for hPPARgamma (MolDock-hPPARgamma)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE4	PPARgamma-act-liver	PPARgamma activation of reporter gene in HepG2 cells (HepG2-RGA-PPARgamma)	Relevant reporter gene assay in human cell line to measure receptor activation	Reliable assay based on positive and negative controls	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
MIE5	FXR-act-liver	Molecular docking model for hFXR (MolDock-hFXR)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE5	FXR-act-liver	FXR activation of reporter gene in HepG2 cells (HepG2-RGA-FXR)	Relevant reporter gene assay in human cell line to measure receptor activation	Reliable assay based on positive and negative controls	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
MIE6	AhR-act-liver	OCHEM QSAR model for AhR (QSAR-OCHEM-AhR-Hepatotoxicity)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping
MIE6	AhR-act-liver	Molecular docking model for hAHR (MolDock-hAHR)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE6	AhR-act-liver	AHR activation of reporter gene in HepG2 cells (HepG2-RGA-AhR)	Relevant reporter gene assay in human cell line to measure receptor activation	Reliable assay based on positive and negative controls	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
MIE7	CAR-act-liver	Molecular docking model for hCAR (MolDock-hCAR)	Relevant to predict binding but does not give information on	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low

			agonism/antagonism			tier RPF data
MIE7	CAR-act-liver	CAR activation of reporter gene in HepG2 cells (HepG2-RGA-CAR)	Relevant reporter gene assay in human cell line to measure receptor activation	Reliable assay based on positive and negative controls	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
MIE8	GR-act-liver	Molecular docking model for hGR (MolDock-hGR)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE8	GR-act-liver	GR activation of reporter gene in HepG2 cells (HepG2-RGA-GR)	Relevant reporter gene assay in human cell line to measure receptor activation	Reliable assay based on positive and negative controls	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
MIE9	RAR-act-liver	Molecular docking model for hRARalpha (MolDock-hRARalpha)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE9	RAR-act-liver	Molecular docking model for hRARbeta (MolDock-hRARbeta)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE9	RAR-act-liver	Molecular docking model for MolDock-hRARgamma (MolDock-hRARgamma)	Relevant to predict binding but does not give information on agonism/antagonism	Assessment not available	Model available, low cost of *in silico* model	Grouping and prioritisation for further testing. Potentially binding energy useful for low tier RPF data
MIE9	RAR-act-liver	RARalpha activation of reporter gene in HepG2 cells (HepG2-RGA-RARalpha)	Relevant reporter gene assay in human cell line to measure receptor activation	Reliable assay based on positive and negative controls	Available low cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
KE1	AOX-decr-liver	ACOX1 (AOX) gene expression in HepaRG cells (HepaRG-PCR-ACOX1)	Relevant gene expression assay human cell line	Reliable assay based on positive and negative controls	Available low-medium cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
KE2	ChREBP-incr-liver	CHREBP (MLXIPL) gene expression in HepaRG cells (HepaRG-PCR-CHREBP)	Relevant gene expression assay human cell line	Reliable assay based on positive and negative	Available low-medium cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data

				controls		
KE3	SREBP-1c-incr-liver	SREBF1 (SREBP-1c) gene expression in HepaRG cells (HepaRG-PCR-SREBF1)	Relevant gene expression assay human cell line	Reliable assay based on positive and negative controls	Available low-medium cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
KE4	FAS-incr-liver	HepaRG-PCR-FASN	Relevant gene expression assay human cell line	Reliable assay based on positive and negative controls	Available low-medium cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
KE5	SCD1-incr-liver	FASN (fatty acid synthase) gene expression in HepaRG cells (HepaRG-PCR-SCD)	Relevant gene expression assay human cell line	Reliable assay based on positive and negative controls	Available low-medium cost assay	Grouping and prioritisation for further testing. Potentially useful for low tier RPF data
KE6	CD36-incr-liver	No assay				
KE7	microsomalbetaox-decr-liver	No assay				
KE8	denovoFA-incr-liver	No assay				
KE9	FAinflux-incr-liver	No assay				
KE10	triglyceride-accum-liver	HCS LipidTOX™ Green neutral lipid stains in HepaRG cells (HepaRG-HCS-triglyceride)	Relevant assay in human cell line to measure triglyceride levels	Reliable assay based on positive and negative controls	Available low-medium cost assay	Grouping and prioritisation for further testing. Useful for *in vitro* RPF data
KE10	triglyceride-accum-liver	quantification of triglyceride accumulation by GC-FID in HepaRG cells (HepaRG-GC-triglyceride)	Relevant assay in human cell line to measure triglyceride levels	Reliable assay based on positive and negative controls	Available low-medium cost assay	Grouping and prioritisation for further testing. Useful for *in vitro* RPF data
KE10	triglyceride-accum-liver	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	Relevant assay in human cell line to measure triglyceride levels	Reliable assay based on positive and negative controls	Available low-medium cost assay	Grouping and prioritisation for further testing. Useful for *in vitro* RPF data
KE11	mitochondrial-disrupt-liver	Mitochondiral disruption measured in the Seahorse XF Cell Mito Stress Test in	Relevant assay in human cell line to measure mitochondrial disruption	Reliable assay based on positive and negative	Available low-medium cost assay	Grouping and prioritisation for further testing. Useful for *in vitro* RPF data

		HepaRG cells (HepaRG-Mito)		controls		
KE12	cytoplasm-displ-liver	No assay				
KE13	nucleus-distort-liver	No assay				
KE14	ER-stress-liver	No assay				
KE15	FattyCells-liver	No assay				
AO	Steatosis-liver	COSMOS QSAR Nuclear Receptor model (QSAR-COSMOS-NR-Hepatotoxicity)	Indirectly relevant since predicts binding to nuclear receptors in the steatosis AOP	Validated for prediction of liver steatosis in EuroMix project using training set. Reasonable reliability.	Model available, low cost of *in silico* model	Grouping
AO	Steatosis-liver	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	Relevant for prediction of liver steatosis	Validated for prediction of liver steatosis in EuroMix project using training set. High reliability.	Model available, low cost of *in silico* model	Grouping
AO	Steatosis-liver	fatty liver cells (histopathology) in rat 28 day study (Rat28day-FattyCells)	Relevant *in vivo* study to measure liver steatosis	Based on OECD test guideline	*In vivo* test available, high cost	Grouping and RPF data

The identified *in silico* models and *in vitro* and *in vivo* assays are described in tables (i.e. Excel sheets) for use in the EuroMix toolbox. The *in vitro* and *in vivo* assays are described as responses measured in test systems. The responses are linked to the corresponding effects (KE). The Excel file with the sheets are uploaded into the EuroMix toolbox. The Excel sheets contain the following information:

Test systems
The Test systems table describes the *in vivo* or *in vitro* system where responses are measured.

Name	**Description**	**Example. Similar data is included for each test system**
idTestSystem	Unique identification code of the test system	HepaRG
Name	Name of the test system	HepaRG
Description	Additional description or label of the test system	Terminally differentiated hepatic cells derived from a human hepatic progenitor cell line
TestSystemType	The type of the test system, i.e., in-vivo, cell-line, etc	CellLine
Organ	If applicable, the organ that the cells originate from associated with the *in vitro* test-system	liver
Species	If applicable, the species associated with the test-system	human
Strain	If applicable, the strain of the species associated with the test-system	
RouteExposure	If applicable, the route of exposure associated with the *in vivo* test-system, oral, dermal, inhalation, s.c., i.v	
GuidelineMethod	Reference to test guideline	
References	External reference(s) to other sources containing more information about the test system. E.g., publications, website, documents	www.heparg.com

Responses table
The Responses table describes the specific responses measured in the test system.

Name	**Description**	**Example. Similar data is included for each response**
idResponse	Unique identification code of the response. In the EuroMix data collection, a EuroMix coding system has been set up in which the id of the test	HepaRG-AdipoRed-72h

	system prefixes the id of the response. E.g., 'HepaRG-PCR-PPARA', 'RatWEC-PCR-CYP26a1' and 'MouseDevelopmental-FacialPrimordia-malformed-E9'.	
Name	Name of the response.	HepaRG-AdipoRed-72h
Description	Additional description or label of the response.	quantification of triglyceride accumulation at 72h using the AdipoRed assay
idTestSystem	Unique identification code of the test system.	HepaRG
ResponseType	The data type of the response measurements (e.g., continuous multiplicative, ordinal, categorical).	ContinuousMultiplicative
ResponseUnit	If the response type is Continuous, then this should be the unit of the response, e.g., kg.	
GuidelineMethod	Reference to the test method guideline, e.g., standaridised assay kit.	

Effect representations table

The Effect representations table describes how the responses are also linked to specific effects.

Name	**Description**	**Example. Similar data is included for effect representation**
idEffect	Identifier of the effect	triglyceride-accum-liver
idResponse	Identifier of the response	HepaRG-AdipoRed-72h
BenchMarkResponse	The Benchmark response value used in benchmark dose modelling	20
BenchMarkResponseType	Specifies how the BenchMarkResponse is expressed, relative to the response at zero dose. E.g percentage change	PercentageChange

QSAR models
The QSAR models table describes the QSAR models and how they are linked to specific effects.

Name	Description	Example. Similar data is included for each QSAR model
idModel	The unique identification code of the QSAR model	QSAR-FERA-Steatosis
Name	The name of the QSAR model	QSAR-FERA-Steatosis
Description	Description of the QSAR model	FERA QSAR model for liver steatosis
idEffect	The effect code	Steatosis-liver
Accuracy	Accuracy of the QSAR model	0,75
Sensitivity	Sensitivity of the QSAR model	0,74
Specificity	Specificity of the QSAR model	0,76
Reference	External reference(s) to sources containing more information about the QSAR model	

Molecular docking models
The molecular docking models table describes the molecular docking models and how they are linked to specific effects.

Name	Description	Example. Similar data is included for each molecular docking model
idMolecularDockingModel	The unique identification code of the molecular docking model	MolDock-hPXR
Name	The name of the molecular docking model	MolDock-hPXR
Description	Description of the molecular docking model	Molecular docking model for hPXR
idEffect	The effect code	PXR-act-liver
Reference	External reference(s) to sources containing more information about the molecular docking model.	

13.4 Annex B4 - Example of grouping substances into assessment groups

This example of grouping substances into assessment groups is based on work performed in the EuroMix project. However, the example is only based on a few substances and not all possible substances that should be considered for grouping. The example is only for illustration purposes and should not be seen as a final assessment group that can be used for mixture risk assessment.

The example follows the methodology described in the EuroMix handbook.

Level of grouping
The grouping is at the level of liver steatosis.

AOP network
The AOP network for liver steatosis developed in the EuroMix project is used for the grouping, see Annex B2 and B3.

Substance category to be assessed
The substances that are considered are a subset of pesticides, named A-I.

Collection of data
In silico and *in vitro* data is from EuroMix and *in vivo* data from an EFSA supporting publication (RIVM, ICPS, ANSES, 2013).

Organising the data
The data was organised into lines of evidence using the table in Annex A5, see Table 1.

Table 1. Data for grouping, organised in lines of evidence.

Substance	Key event in the AOP network (organised according to MIE, intermediate KEs, AO)	Study type (organised according to *in silico*, *in vitro*, *in vivo* data, human study)	Assay (specific assay used)	Main study result (e.g. positive, negative, for positive assays/studies BMD/ NOAEL is given for the specific effect measured in the assay/study)	Reliability of data (low, medium, high)	Relevance of data (low, medium, high)
A	KE 10: triglyceride-accum-liver	*In vitro*	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	750 µM (BMD-50)	high	medium
A	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	negative	high	medium
A	AO: Steatosis-liver	*In vivo*	Rat 28 day study-fatty changes	120 mg/kg bw day (NOAEL)	high	high
B	KE 10: triglyceride-accum-liver	*In vitro*	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	no data		
B	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	positive	high	medium
B	AO: Steatosis-liver	*in vivo*	Rat 2 year study-fatty changes	1 mg/kg bw day (NOAEL)	high	high
C	KE 10: triglyceride-accum-liver	*In vitro*	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	6 µM (BMD-50)	high	medium
C	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	positive	high	medium
C	AO: Steatosis-liver	*In vivo*	Mouse 90 day study-fatty changes	34 mg/kg bw day (NOAEL)	high	high
D	KE 10: triglyceride-	*In vitro*	quantification of triglyceride	no data		

	accum-liver		accumulation using the AdipoRed assay (HepaRG-AdipoRed)			
D	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	positive	high	medium
D	AO: Steatosis-liver	*in vivo*	Dog 1 year study-fatty changes	0.025 mg/kg bw day (NOAEL)	high	high
E	KE 10: triglyceride-accum-liver	*In vitro*	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	19 µM (BMD-50)	high	medium
E	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	negative	high	medium
E	AO: Steatosis-liver	*In vivo*	Rat 2 year study-fatty changes	0.53 mg/kg bw day (NOAEL)	high	high
F	KE 10: triglyceride-accum-liver	*In vitro*	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	24 µM (BMD-50)	high	medium
F	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	positive	high	medium
F	AO: Steatosis-liver	*in vivo*	Rat 90 day-fatty changes	4 mg/kg bw day (NOAEL)	high	high
G	KE 10: triglyceride-accum-liver	*In vitro*	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	610 µM (BMD-50)	high	medium
G	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	negative	high	medium
G	AO: Steatosis-liver	*In vivo*	Mouse carcinogenicity study-fatty changes	115 mg/kg bw day (NOAEL)	high	high
H	KE 10: triglyceride-accum-liver	*In vitro*	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	250 µM (BMD-50)	high	medium

H	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	negative	high	medium
H	AO: Steatosis-liver	*in vivo*	Mouse 2 year study-fatty changes	5.7 mg/kg bw day (NOAEL)	high	high
I	KE 10: triglyceride-accum-liver	*In vitro*	quantification of triglyceride accumulation using the AdipoRed assay (HepaRG-AdipoRed)	3000 µM (BMD-50)	high	medium
I	AO: Steatosis-liver	*In silico*	FERA QSAR model for liver steatosis (QSAR-FERA-Steatosis)	negative	high	medium
I	AO: Steatosis-liver	*in vivo*	Mouse 90 day study-fatty changes	543 mg/kg bw day (NOAEL)	high	high

Assessing the data for reliability and relevance

The data was assessed for reliability and relevance and the result was reported in Table 1. All data were given the score “high” reliability since they were generated from an *in silico* model with high accuracy, an *in vitro* assay performed according to standard operating procedures, and *in vivo* tests performed according to OECD test guidelines. *In silico* data were given the score “medium” relevance since they were generated from a QSAR model, the *in vitro* data were given the score “medium” relevance since they were generated from an *in vitro* assay measuring an intermediate KE and the *in vivo* data were given the score high relevance since they measured the AO *in vivo*.

Decision of group membership

The decision on group membership was based on the information in Table 1 and all substances were included in the assessment group since they had adequate data both for KE 10 and the AO.

Reporting of group membership

The assessment group membership was reported as 1 (included) in the AssessmentGroupMembership table for use in the EuroMix toolbox, see figure 1.

	A	B	C	D	E
1	idGroupMembershipModel	idEffect	idSubstance	GroupMembership	
2	1	Steatosis-liver	idSubstance-A	1	
3	1	Steatosis-liver	idSubstance-B	1	
4	1	Steatosis-liver	idSubstance-C	1	
5	1	Steatosis-liver	idSubstance-D	1	
6	1	Steatosis-liver	idSubstance-E	1	
7	1	Steatosis-liver	idSubstance-F	1	
8	1	Steatosis-liver	idSubstance-G	1	
9	1	Steatosis-liver	idSubstance-H	1	
10	1	Steatosis-liver	idSubstance-I	1	
11					

AssessmentGroupMembershipModels | AssessmentGroupMemberships

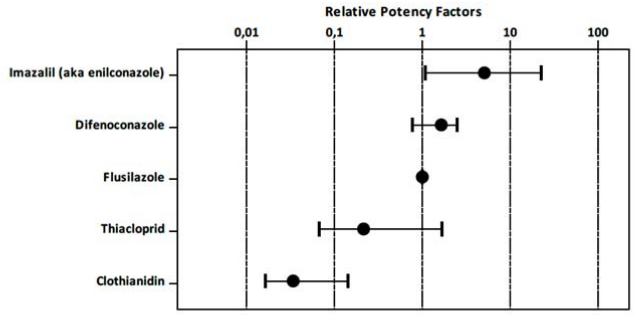
Figure 1. Table AssessmentGroupMembership reporting the results of the grouping of substances into the assessment group Steatosis-liver for use in the EuroMix toolbox.

13.5 Annex B5 - Example of calculation of RPFs from in vitro data using benchmark dose method

This example of calculation of RPFs based on *in vitro* data using benchmark dose method is based on work performed in the EuroMix project. However, the example is only for illustration purposes and the results should not be seen as final RPFs that can be used for mixture risk assessment. The example follows the methodology described in the EuroMix handbook.

The example is based on five substances that have been grouped in an assessment group for liver steatosis. Please note that most real assessment groups contain many more substances. Accumulation of triglycerides in HepaRG cells was measured using AdipoRed assay after 72 hour exposure to the substance. Accumulation of triglycerides is a KE in the AOP network leading to liver steatosis. Dose response modelling and calculation of RPFs was done in the EuroMix toolbox. Information was provided in Excel sheets for the following modules in the EuroMix toolbox: AOP networks (liver steatosis), Effect relations (connecting the effects in the AOP network), Test systems (HepaRG), Responses (HepaRG-Adipored-72h), Effects (steatosis-liver), Effect representations (information that the response is used to measure the effect, and the Benchmark response used), Substances and Dose-response experiments.

In this example the purpose was to derive BMDs to calculate RPFs for a group of substances, but the BMDs will not be used as a POD for the risk assessment. In this case it is not necessary to select a BMR that reflects a no effect level, but a higher BMR can be chosen that is statistically more robust, and anywhere between the minimal and maximal observed responses, provided the dose-response curves are parallel. In this example the BMR of 20% was used that was suitable for the *in vitro* datasets. Flusilazole was selected as index substance based on the criteria described in the handbook. The benchmark dose software Proast is integrated in the EuroMix toolbox allowing benchmark dose analysis of dose response data to be done within the toolbox. Dose response models were fitted to the dose response data. The BMD from the models were used to calculate PODs and the RPFs were calculated based on the PODs. Uncertainty analysis was included using resampling of the PODs/RPFs. The results are shown in Figure 1.



Substance name	Substance code	RPF (nominal) scaled to reference	Relative potency factor lower bound (2.5 percentile)	Relative potency factor upper bound (97.5 percentile)
Clothianidin	RF-0101-001-PPP	0.0338	0.0163	0.142
Difenoconazole	RF-0133-001-PPP	1.64	0.769	2.48
Flusilazole	RF-0218-001-PPP	1	1	1
Imazalil (aka enilconazole)	RF-0246-001-PPP	5.08	1.08	22.7
Thiacloprid	RF-0417-001-PPP	0.214	0.0669	1.67

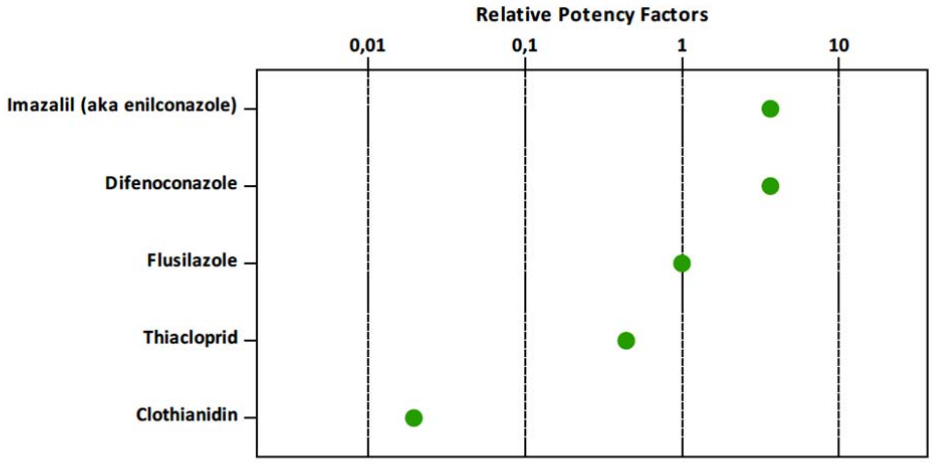
Figure 1. Example of calculation of RPFs based in *in vitro* data using the EuroMix toolbox. Flusilazole was chosen as index substance.

13.6 Annex B6 - Example of imputation of POD using Munro collection of TTC

This example of imputation of PODs using the Munro collection of TTCs and calculation of RPFs based on the imputed PODs is based on work performed in the EuroMix project. However, the example is only for illustration purposes and the results should not be seen as final RPFs that can be used for mixture risk assessment. The example follows the methodology described in the EuroMix handbook.

The example is based on five substances that have been grouped into an assessment group for liver steatosis. Please note that most real assessment groups contain many more substances. In this fictional example, NOAELs for liver steatosis were derived from *in vivo* studies for three of the substances but two substances lacked PODs. RPFs were calculated using the EuroMix toolbox and based on the available *in vivo* NOAELs and PODs imputed based on the 5th percentile of the NOAELs in the Munro collection for Cramer class 3. Cramer class 3 was used since the two substances with missing PODs belong to that class. Information was provided in Excel sheets for the following modules in the EuroMix toolbox: AOP networks (liver steatosis), Effect relations (connecting the effects in the AOP network), Effects (steatosis-liver), Point of departures (NOAELs for three of the substances) and Substances. Flusilazole was selected as index substance based on the criteria described in the handbook.

PODs were imputed for imazalil and difenoconazole and *in vivo* NOAELs were used for flusilazole, thiacloprid and clothianidin. The imputed PODs are conservative using the 5th percentile of NOAELs from Cramer class 3 and result in the same RPF of 3.66, see Figure 1.



Substance name	Substance code	RPF (nominal) scaled to reference
Clothianidin	RF-0101-001-PPP	0.0196
Difenoconazole	RF-0133-001-PPP	3.66
Flusilazole	RF-0218-001-PPP	1
Imazalil (aka enilconazole)	RF-0246-001-PPP	3.66
Thiacloprid	RF-0417-001-PPP	0.442

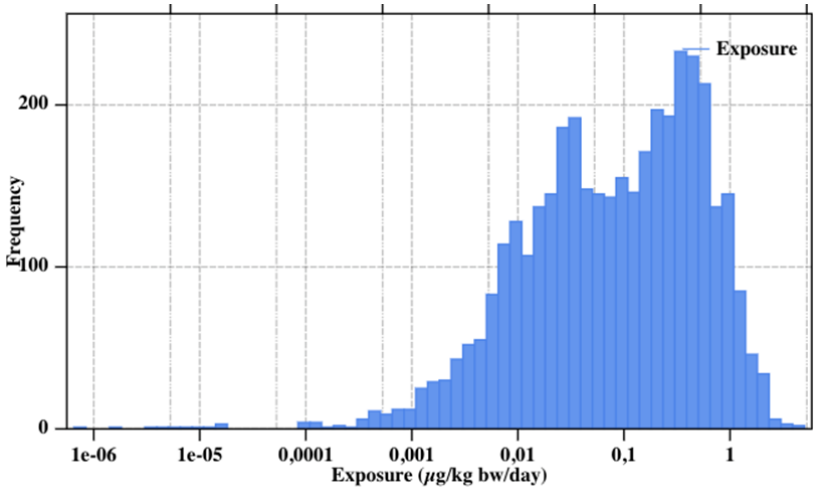
Figure 1. Example of imputation of missing PODs and calculation of RPFs using the EuroMix toolbox. PODs were imputed for imazalil and difenoconazole. Flusilazole was chosen as index substance.

13.7 Annex B7 - Example of risk characterization using dose addition

This example of how the potency-scaled combined exposure can be modelled in the EuroMix toolbox and expressed is based on work performed in the EuroMix project. However, the example is only for illustration purposes and the results should not be seen as exposure assessment that can be used for mixture risk assessment. The example follows the methodology described in the EuroMix handbook.

The combined exposure to 144 pesticides in the assessment group for liver steatosis was modelled probabilistically in the EuroMix toolbox. This example used food consumption data from a food consumption survey, concentration data from measurements of levels of substances in food items, food recipes for food conversion and processing factors. RPFs were based on *in vivo* NOAELs for liver steatosis. Flusilazole was used as index substance.

The potency-scaled combined exposure is expressed as a distribution and compared to the POD of the index substance. Different percentiles of the exposure distribution and the corresponding MOEs can be calculated, Figure 1



Percentage	Exposure (µg/kg bw/day)	Percentage of PoD (%)	Nominal margin of exposure
50.00	0.09611	0.02	5514
90.00	0.7403	0.14	716
95.00	1.059	0.20	500.6
99.00	1.872	0.35	283.1
99.90	3.03	0.57	174.9
99.99	4.401	0.83	120.4

Figure 1. Example of distribution of the potency-scaled combined dietary exposure calculated using the EuroMix toolbox. The exposure distribution can be compared to the POD of the index substance, in this example 530 µg/kg bw per day. The table shows different percentiles of the exposure distribution and the corresponding MOE.

13.8 Annex B8 - Example of uncertainty analysis

The example of uncertainty analysis for mixture risk assessment is based on work performed in the EuroMix project during development of the EuroMix methodology and tools, but should not be seen a real mixture risk assessment.

Aspect	**Identified uncertainties**	**Analysis of uncertainty**
Grouping of substances into assessment groups	Grouping of pesticides based on liver steatosis as common effect/adverse outcome using in vivo, in vitro and in silico data. Pesticides that should be included in the group might be missing due to missing or uncertain toxicity data. Pesticides that do not cause steatosis might be included in the group also due to missing or uncertain toxicity data.	Uncertainty assessed qualitatively or quantitatively using probabilistic resampling of probabilities for group membership.
Choice of toxicity data to derive POD	In vivo toxicity studies for the effect liver steatosis in rat were used. Uncertainty related to the data collection and selection of the studies.	Uncertainty assessed qualitatively.
Calculation of RPFs	NOAEL used for most substances and LOAEL for the ones lacking a NOAEL. Uncertainty related to use of NOAEL/LOAELs rather than BMD.	Uncertainty assessed qualitatively.
Extrapolation between *in vitro* and *in vivo* studies	Not applicable. Only in vivo studies used for RPFs.	
Lack of toxicity data	Not applicable. In vivo data available for all substances.	
Consumption data	Food consumption survey for the adult population in Netherlands, using 24 h-recall on 2 non-consecutive days. Uncertainty related to the representability of the study and the method used.	Sampling uncertainty quantitatively assessed using the bootstrap method.
Concentration data	Concentration data from European control and monitoring programmes, using objective or selective sampling. Uncertainty related to the sampling and analytical methods.	Sampling uncertainty quantitatively assessed using the bootstrap method.
Non-detects, concentration measurements below the limit of detection	Non-detects in 99.3% of the measurements. Zero value was used for non-detects resulting in uncertainty of the true values.	Uncertainty assessed qualitatively.
Lack of concentration data	For 17 substances no concentration data was available and legal residue limits were used resulting in uncertainty in true values.	Uncertainty assessed qualitatively.
Conversion of food-as-eaten to	Dutch recipes used for conversion.	Uncertainty assessed

food-as-measured and processing factors	Processing factors available for 46 of 144 substances resulting in uncertainty in true concentration levels. Conversion between parent compound and residue definitions resulting in uncertainty in true concentration levels.	qualitatively.
Other (non-dietary) routes of exposure	Not applicable. Only dietary exposure.	
Use of the dose addition model	Dose addition used as conservative default.	Uncertainty assessed qualitatively
Possible interactions (synergism or antagonism)	The available information does not suggest that interactions take place.	Uncertainty assessed qualitatively

14 Annexes C - Training material for the EuroMix toolbox

14.1 Annex C1 - Training material describing how to group substances based on *in silico* data in the EuroMix toolbox

The training material describes a step-by-step approach how to calculate membership in the assessment group using a list of included/excluded substances based on QSAR models. The training material is included in deliverable 10.6.

14.2 Annex C2 - Training material for calculation of relative potency factors using the module Relative potency factors in the EuroMix toolbox

The training material describes a step-by-step approach how calculate RPFs based on dose-response data using benchmark dose method. The training material is included in deliverable 10.6.

14.3 Annex C3 - Training material for imputation of missing POD data in the EuroMix toolbox

The training material describes a step-by-step approach how to impute missing POD data from Munro collection of TTCs. The training material is included in deliverable 10.6.

14.4 Annex C4 - Training material for performing a probabilistic dietary exposure assessment of mixtures in the EuroMix toolbox

The training material describes a step-by-step approach how to perform a probabilistic dietary exposure assessment of mixtures in the EuroMix toolbox. The training material is included in deliverable 10.6.

14.5 Annex C5 - Training material for performing exposure assessment using a combination of dietary and non-dietary sources in the EuroMix toolbox

The training material describes a step-by-step approach how to perform exposure assessment using a combination of dietary and non-dietary sources in the EuroMix toolbox. The training material is included in deliverable 10.6.