1

2

Output metadata	
Output category	Guidance
Date adopted [Panel/SC]	
Date approved [non-Panel/SC]	
Author list [names or ORCID in order of authorship WIN, beginning with EFSA Panel]	EFSA Working Group
DOI	10.2903/j.efsa.201Y.xxxx10.2903/j.efsa.201Y.xxxx
Requestor	[European Commission]
Output number	ON1234ON1234
Question number(s) [separate multiple entries with commas]	EFSA-Q-201Y-XXXXX
Correspondence	pesticidespeerreview@efsa.europa.eu
Short title [header]	
Panel members	
Copyright for non-EFSA content	
Acknowledgments	
Minority opinion	
Competing interests	
Waiver	
Note	
Legal notice	
Disclaimer	

Example data are in square brackets and need to be updated or deleted before dispatch to Wiley

Revised guidance on the risk assessment of plant protection products on bees (*Apis mellifera*, *Bombus* spp. and solitary bees)

European Food Safety Authority (EFSA)

Abstract

The European Commission asked EFSA to revise the Guidance on the risk assessment for honey bees, bumble bees and solitary bees. This Guidance described how to perform risk assessment for bees from plant protection products, in accordance with Regulation (EU) 1107/2009. The Guidance Document is a review of EFSA's existing guidance document which was published in 2013. The Guidance Document outlines tiered approach for exposure estimation in different scenarios and tiers. It includes hazard characterisation and provides risk assessment methodology covering dietary and contact exposure. The document provides also recommendations for higher tier studies, risk from metabolites and mixture.

Keywords

Bees, pesticides, risk assessment, higher tier studies

Summary

In 2019, EFSA received a mandate to revise the 2013 bee guidance with several Terms of Reference (ToRs) including collecting data on bee mortality, revise the requirements for field studies, revise the crop attractiveness for pollen and nectar and risk assessment methodologies and support the definition of specific protection goals (SPGs).

In line with the mandate, EFSA has carried out an evidence-based revision based on systematic approaches for several aspects, consulted stakeholders and Member States during the process and supported risk managers (RM) to define the magnitude dimension of the SPG for honey bees, and discuss the approach for bumble bees and solitary bees. For honey bees, RM agreed a value of 10% as the maximum permitted level of colony size reduction. For bumble bees and solitary bees, RM did not define a quantitative magnitude of acceptable effects due to the lack of data. However, there was a general consensus to require more frequently higher tier studies in order to gain more robust data on the effects of pesticides on those bees allowing their better protection.

This guidance document allows to address the risk to honey bees that are exposed to plants protection goals (PPPs) in agricultural areas, by following a tiered approach for the exposure estimation and the effect assessment. For bumble bees and solitary bees, the guidance allows to tailor the studies that need to be generated.

The guidance considers the exposure via contact when bees are oversprayed and/or the exposure via diet when bees consume contaminated pollen and nectar in different exposure scenarios, that include intentionally treated areas and accidentally contaminated surrounding areas.

Since both exposure and effect assessments are operationalised in the tiered approach, an exposure-Tier and an effect-Tier have been defined. In the exposure tiers, residue intake or residue deposition need to be quantified by calculating the Predicted Exposure Quantity (PEQ) to address the dietary and the contact route of exposure of the bees following the use of a PPP in the field. In the effect tiers the imposed exposure is called 'Dose' in the laboratory tests or 'Estimated Exposure Dose' in the higher tier tests.

For both exposure and effect assessment sides, the bee routes of exposure are addressed by considering the different timescale of effects (acute and chronic) and the different life stages (adults and larvae). To this purpose in this guidance document four risk cases have been defined: 1) Acute-contact; 2) Acute-dietary; 3) Chronic-dietary, 4) Larvae-dietary.

The exposure estimation in the different Tiers will provide PEQ for each of the above risk cases and it is indicated as PEQ_j where the suffix j indicates the four risk cases.

Mathematical models to estimate the PEQ_j in the different tiers, have been revised and reparametrized, including systematic literature reviews for the key parameters related for example to a better estimation of the food consumption. Guidance had been developed for appropriate refinement options for many of the parameters (Tier 2 exposure assessment).

In parallel, the effect assessment will provide median lethal response dose $LD50_j$ and $slope_j$ (the parameter describing the steepness of the dose-response relationship obtained from standard laboratory tests). These two parameters are used to connotate the dose-response relationship for each risk case. Methods for extrapolating $LD50_j$ values between bee species have been developed.

As part of the effect-tier assessment of the PPP, the guidance document suggests addressing two additional aspects for honey bees: the potential for the compound under evaluation for showing increasing toxic effects due to long-term exposure to low doses (Time Reinforced Toxicity assessment (TRT) and potential concerns due to sublethal effects.

The higher effect tier is formed by different type of studies e.g. semi-field, colony feeder and/or field tests. The field tests represent the highest level of experimentally feasible effect assessments on bees foraging at local and larger scales (treated field, immediate off-field areas and possibly landscape). In principle, higher tier effect studies can be supported by population-level modelling of effects for different ecological and agricultural practice scenarios, but such models first would need to be developed, calibrated and evaluated for their use in regulatory bee risk assessment.

Finally, the guidance document provides a methodology and risk assessment scheme for metabolites, mixtures and a consideration of possible risk mitigation measures.

Table of Contents

1 Introduction

1.1 Background and Terms of Reference as provided by the European Commission

In 2013, EFSA issued a guidance document on the risk assessment (RA) of plant protection products on bees (*Apis mellifera*, *Bombus spp.* and solitary bees) (EFSA, 2013), which has not been fully implemented in the regulatory framework owing to some lack of consensus triggering a request for revision by the Member States.

In March 2019 the European Commission mandated EFSA to review the guidance, by including in the mandate, the following Terms of References (ToRs):

1. *take account of the feedback from Member States and stakeholders on the EFSA (2013) guidance document* (**ToR1**)*;*
2. *provide a review and summary of the evidence as regards bee background mortality, in particular considering realistic beekeeping management for Apis mellifera and natural background mortality. EFSA is requested to provide this summary in a separate document from the guidance document* (**ToR2**)*;*
3. *review the list of bee-attractive crops in particular considering presence of bees, guttation and agricultural practices (harvesting time before or after flowering). This reviewed list shall also mention at which growing phases (e.g. BBCH codes) a crop is considered bee-attractive* (**ToR3**)*;*
4. *review the current risk assessment methodologies in light of recent scientific research and developments e.g. exposure estimation, relevance of the exposure scenarios (e.g. weed scenario) and relevance of some risk assessment schemes. Available relevant guidance developed by Member States should be considered (e.g. draft guidance document on seed treatments and/or its follow up work)* (**ToR4**)*;*
5. *review the requirements for higher tier testing, in particular by reconsidering the magnitude of detectable effects vs the statistical power and validated population modelling in light of realistic agro-environmental conditions* (**ToR5**)*;*
6. *take into account planned and on-going discussions initiated by the Commission on defining specific environmental protection goals and review the risk assessment guidance based on the specific protection goals agreed during this process* (**ToR6**)*.*

To address the **ToR1**, EFSA established an *ad hoc* group of stakeholders that was consulted during the review process in parallel with Member States (MSs). The consultations performed were used to tailor the review and to select the most appropriate methodological approaches[1].

The **ToR2** was addressed by collecting data on background mortality of bees with a systematic literature search and the details have been reported in a standalone document (EFSA et al., 2020).

To address the **ToR3** and **ToR4**, the working group (WG) developed a protocol which is available in the Annex A of the Supplementary document of this guidance. In this Supplementary document, the WG reported detailed explanation, data, results of the

[1] https://www.efsa.europa.eu/en/topics/topic/pesticides-and-bees-guidance-review

preparatory work that was used as basis to review the EFSA guidance 2013. For the revision of the requirements of higher tier studies (**ToR5**) the WG considered both the exposure and the effect assessment in light of the Specific Protection Goals (SPGs) agreed by risk managers (RMs).

For **ToR6,** the WG provided support to RMs for decision making process on (SPG), by considering the ongoing activities of the European Commission on this topic and the feedback received by RMs (EFSA et al., 2021).

1.2 Legal framework

Among its approval criteria, Regulation (EC) No. 1107/2009 states that plant protection products (PPPs) may be approved only if they have no unacceptable effects on the environment, including non-target species, and impact on biodiversity and the ecosystems. In addition to this, Regulation (EC) 1107/2009[2] in its Annex II, point 3.8.3 gives an explicit approval criterion for honey bees:

An active substance, safener or synergist shall be approved only if it is established following an appropriate risk assessment on the basis of Community or internationally agreed test guidelines, that the use under the proposed conditions of use of plant protection products containing this active substance, safener or synergist:

- *will result in a negligible exposure of honey bees, or*
- *has no unacceptable acute or chronic effects on colony survival and development, taking into account effects on honey bee larvae and honey bee behaviour.*

Regulation (EC) No. 283/2013[3] and Regulation (EC) No. 284/2013[4] set the specific data requirements for the active substances and the PPPs, respectively, and give indication of the standard agreed study protocols that are applied.

As 'unacceptable effects' is not further defined in the Regulation, this constitutes a very generic protection goal which needs to be translated into specific (operational) protection goals (SPGs) that can be linked in a transparent way to selected risk assessment schemes in guidance documents. EFSA PPR Panel (2010) and EFSA Scientific Committee (2016) proposed a methodology to define SPGs based on ecosystem services. The underlying principle of this methodology is that the general protection goal of the legislation is achieved by minimising the impact on the ecosystem services via the protection of their providers (i.e. services providing units). This methodology has been used in the previous EFSA guidance document (EFSA, 2013) and proposed by the European Commission for the definition of environmental

[2] Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJ L 309, 24.11.2009, p. 1–50.

[3] Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, p. 1–84.

[4] Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, p. 85–152.

SPGs for PPPs (EFSA et al., 2021). Therefore, the definition of the SPG agreed by RMs has been still based on this method (see Section 3).

1.3 Bee ecology

The target groups for this guidance document are honey bees (*Apis mellifera*), bumble bees (*Bombus* spp.) and solitary bees. These three groups have differing life histories, so in this section we include a short summary of the major differences between these groups.

1.3.1 General information on bee life history

Bee species can be classified by their social organisation and can be social (9.4%), solitary (77.4%), or parasitic (13.2%) (Danforth et al., 2019). Eusociality is characterised by a social colonial existence, where adult colony members belong to two or more overlapping generations, care cooperatively for the young, and are divided into reproductive and nonreproductive (or at least less-reproductive) castes (Wilson and Holldobler, 2005, Danforth et al., 2019). Both honey bees and bumble bees are eusocial and produce three types of adult bee: workers, drones, and queens. Honey bees have perennial colonies while social bumble bees have annual colonies where the founding queen also has a solitary phase. During the social phase, the workers are the most abundant type of adult and consist of female bees which will usually not reproduce but will forage for food and maintain the nest. Drones and queens are reproductive individuals. Drones are males that leave the colony to reproduce and the queen is a reproductive female that establishes or allows a colony to develop over time. Solitary bees do not form colonies and lack the worker caste; the adults are simply males and females and all females can reproduce (Danforth et al., 2019). Parasitic taxa often rely on closely related host species (Danforth et al., 2019), which is why they are here treated as part of their host bee group.

Bees can collect various resources from plants such as pollen, nectar, and less commonly, oil or perfumes; some materials used for nesting like resin, soil, and pieces of leaves and petals are also collected (Wcislo and Cane, 1996, Michener, 2000). Various foraging strategies have been described for bees mainly based upon the range of pollen collection from host plant(s). Bees collect pollen and nectar as a food source for their larvae and themselves and, in doing so, pollinate the flowers of the plants they forage from. This mutual dependency of plants on animal pollination and bees on pollen foraging is the result of co-evolution.

1.3.2 Honey bees (*Apis mellifera*)

Honey bees are eusocial bees that live in large (thousands to tens of thousands of individuals) perennial colonies with a single egg-laying queen. In Europe, honey bees are primarily the western honey bee (*Apis mellifera*) which is represented by 11 European subspecies and several locally adapted populations referred to as ecotypes (Strange et al., 2007, Meixner et al., 2013). Whilst non-managed, feral, colonies exist, honey bees in Europe are primarily a managed species raised in artificial nests, i.e., hives. Honey bee nests consist of wax structures made of hexagonal cells, known as combs, that are used to store food and rear young. Honey bees have a highly structured social system. The queen specialises in laying eggs, whilst workers exhibit temporal polyethism, which means tasks are allocated by age; younger workers tend the developing young and maintain the colony whilst older workers forage for pollen, nectar, honey dew, plant resin and water (Seeley, 1982) Workers are central place foragers and obtain food in an area around the hive, foraging at an average distance of 1.6 km with a 90th percentile range of 3.8 km. Honey bee nests contain large reserves of food,

which allows the colony to persist throughout the winter and through sustained periods of poor weather (Seeley, 2009).

1.3.3 Bumble bees (*Bombus* spp.)

Bumble bees are mostly eusocial bees that form small (tens to hundreds of individuals) annual colonies with a single egg-laying queen although some species within the genus can parasitise on the eusocial species. Globally, there are approximately 250 species, and 68 occur in Europe (Nieto et al., 2014). Some bumble bee species, primarily *B. terrestris* in Europe, are used as managed pollinators but most species are wild. Bumble bees also construct their nest structure from wax but lack the regular structure and appearance of honey bee nests. Unlike honey bees, each queen is responsible for establishing a colony, laying eggs, and foraging for pollen and nectar until the first generations of workers develop. After the first generation of larvae become adults, the queen remains in nest to lay eggs. Whilst there is limited evidence that bumble bee workers exhibit alloethism ((Goulson et al., 2002, Foster et al., 2004, Gardner et al., 2007), where tasks are allocated by body size, with smaller workers tending the developing young and maintaining the colony whilst larger workers forage for pollen and nectar, these divisions are not consistent enough to readily distinguish between in-hive and forager workers based on size alone. As bumble bee colonies do not persist over winter the colony stores much less food than honey bees, usually only enough food to allow the colony to persist through short periods of poor weather (Heinrich, 2004). Like honeybees, bumblebees are central place foragers that obtain food from the area surrounding their colonies. However, given the wider diversity amongst bumble bee species and less understanding of their biology and ecology relative to honey bees it is more difficult to estimate their foraging typical ranges. For example, one study from the UK on four common species estimated a wide range of maximum foraging distances (450 – 758 m) which were attributed to significant differences in fundamental aspects of their ecology (Knight et al., 2005). If the colony is successful, new queens and drones are produced towards the end of the colony cycle. For most species the new queens hibernate during the winter and start new colonies the following spring whilst original colony and the workers and drones will die.

1.3.4 Solitary bees (multiple genera)

Solitary bees are a taxonomically diverse group, of the more than 20,000 described bee species in the world, most are solitary bees (IPBES, 2016), so it is difficult to generalise across species. However, in contrast to honey bees and bumble bees this group is not eusocial, but they are as the other bees central place foragers. There are approximately 1900 solitary bee species in Europe, and they exhibit a diverse range of life-history traits; they may be univoltine (single generation a year) or multivoltine (multiple generations a year), excavate nests in the ground or nest in pre-existing cavities above ground, use a wide variety of nesting substrates (soil, wood, masonry, leaf, and other vegetation), generally, provision their nests only once and producing a relatively small (10's) number of offspring (Danforth et al., 2019). Whilst it is generally accepted that larger solitary bees can fly further distances, is not possible to provide a meaningful average distance for the foraging ranges in solitary bees. For example, a review by (Zurbuchen et al., 2010) reports the maximum foraging distances of different solitary bee species at between 100 m and 6 km. Only a small number of solitary bee species are managed, and *Osmia bicornis* and *O. cornuta* are most commonly used in Europe (IPBES, 2016) so the vast majority of the solitary bees are wild.

1.4 Pathways of PPP exposure for bees

The use of PPPs on the agricultural land may result in the exposure of bees present in both treated and surrounding areas. The exposure may occur via different pathways, depending on the application methods and crop growing systems, as well as the ecology of the bees. A comprehensive analysis and overview of the bee exposure pathways is reported in the EFSA PPR opinion of 2012 (EFSA PPR Panel, 2012).

An overview of the pathways of exposure for bees to PPPs which are evaluated in this Guidance Document is shown in Figure 1.

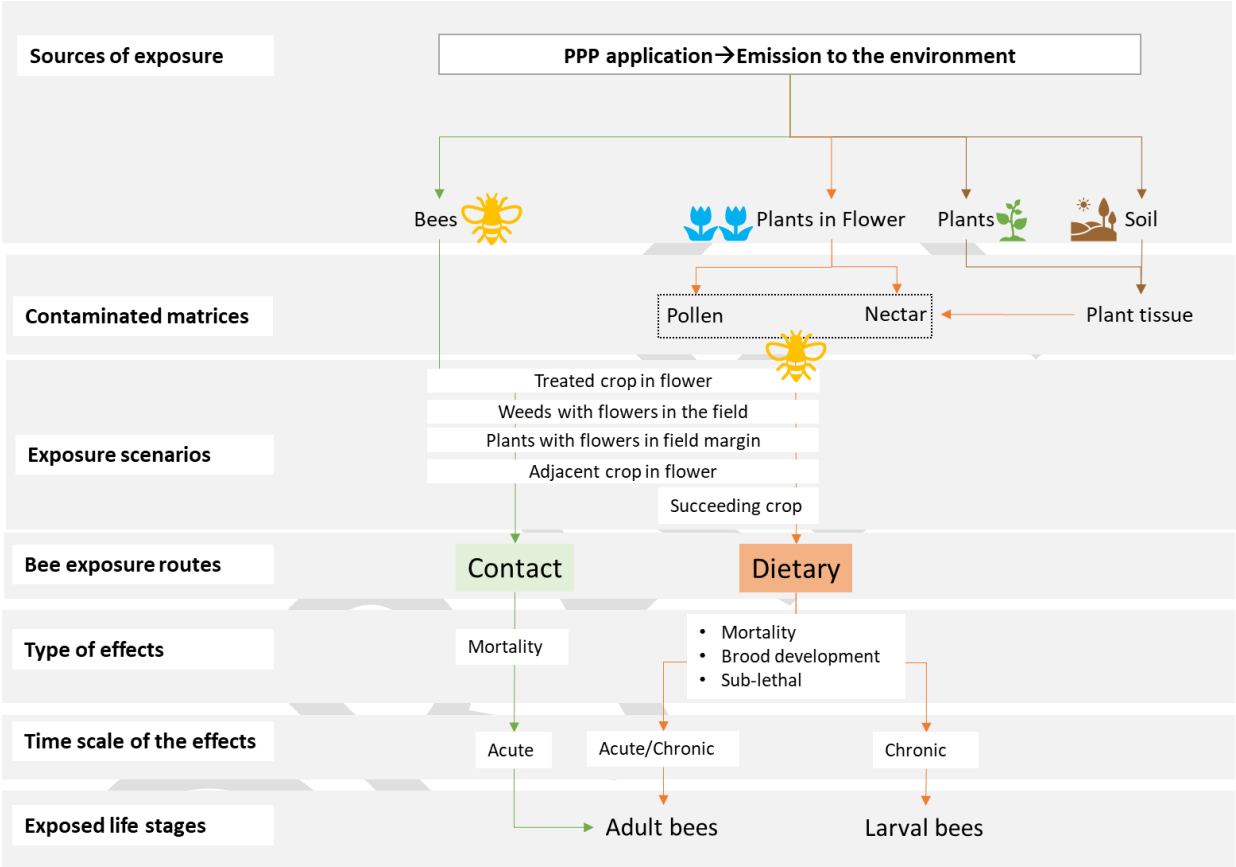


Figure 1: Bee exposure pathways and possible effects in time. Exposure via contaminated matrices other than pollen and nectar or exposure via inhalation are not included.

The methods of application of PPPs together with the crop growing systems determine the emission of a PPP in the environment and the source of exposure for bees, e.g. via overspray, dust/spray drift, soil contamination. Bees may come into direct contact with PPPs (e.g., droplets of spray and solid particles in air), or be exposed via contaminated matrices (e.g., by consuming nectar, pollen).

The bee matrices may be contaminated directly (e.g. by spray liquid or dust deposits to the pollen/nectar) or via a series of processes: For example, 1) a PPP is sprayed onto the plant surface (e.g. leaves) → the PPP enters the plant and is distributed through the plant tissue → reaches the reproductive organ(s) → excreted into e.g. pollen and nectar; 2) a proportion of the PPP is sprayed onto the soil → PPP is taken up by the roots of the plant → distributes within the plant → reaches the reproductive organ(s) → diffused to e.g. pollen and nectar.

Bees may encounter PPPs or contaminated matrices either in the areas treated with the PPP (in-crop) or in areas which were not directly treated but have been unintentionally contaminated (off-crop areas). In order to describe the agricultural areas where bees may be exposed to PPPs while collecting pollen and/or nectar and/or because they carry back contaminated pollen/nectar to the colony/nest, various exposure scenarios have been defined:

- Treated crop in flower;
- Weeds in the field unintentionally contaminated or intentionally treated;
- Plants with flowers in the field margin unintentionally contaminated;
- Adjacent crop in flower: unintentionally contaminated;
- Succeeding crop unintentionally contaminated through the soil.

The way the PPPs or their residues reach the bees (different life stages) determines the bee exposure routes, which could be:

- By contact: it occurs when bees enter in physical contact with the PPPs or with contaminated matrices, but that does not involve ingestion.
- By dietary: it occurs when bees orally consume contaminated material and therefore, they ingest residues of PPP with their diet

In this guidance only contact exposure as a result of overspray or dust/spray drift on adult bees is considered, while exposure in situations such as a bee walking repetitively on contaminated surfaces is not considered.

It is acknowledged that for most species of bumble bees and solitary bees nesting in the soil, exposure by contact with contaminated soil/mud/leaves may be relevant as already reported in the EFSA PPR Panel (2012). However insufficient information is still currently available to address these exposures (Gradish et al., 2018, Sgolastra et al., 2019).

Exposure through inhalation is not routinely addressed, as this is considered only a very minor exposure route for the majority of the substances. However, for very volatile substances (e.g., soil fumigants), this exposure route could be relevant, therefore, on a case-by-case basis, it should be addressed (see Section 4.2).

For the dietary, exposure via nectar and pollen for both adult bees and larvae is addressed. Other contaminated matrices (e.g., honey dew, extrafloral nectaries, resin etc.) that could lead to oral residue intake are not explicitly covered in this guidance due to lack of sufficient data to understand their relevance.

The WG has evaluated the relevance of the exposure via consumption of contaminated water by considering the possibility to quantify the water consumption and the frequency and magnitude of water collection. However, data were not sufficient to achieve a reliable estimation of either aspect. In consideration of this, the WG has taken the decision not to include exposure from this contaminated matrix in the risk assessment. More details and data analysis concerning this decision are included in Annex E of the Supplementary document.

The WG recommends that the areas which are not covered in this guidance are addressed in a future revision of the guidance document and recognises the need to generate further research and data. Overall, addressing the knowledge gaps in relation to above bee exposure routes and to other matrices is pivotal for complementing in future risk assessment of bees.

2 Scope of the Guidance document

This document is intended to provide guidance to applicants and risk assessors for the risk assessment of bees in the context of the evaluation of PPPs and their active substances under Regulation (EC) 1107/2009 for authorisation process at Member State level and the approval at EU level, respectively. In particular, this guidance covers risks to honey bees, bumble bees and solitary bees that may occur when directly exposed to PPPs. Effects caused indirectly by the use of PPPs such as, for example, removal of flowering weed by herbicides, or direct (and indirect) effects on other insect pollinators are out of its scope.

The guidance document covers the risk assessment for chemical active substance, mainly applied as spray, seed treatment and granules, although the principles of proposed risk assessment schemes are relevant for other methods of applications. The guidance document does not cover the risk assessment for micro-organism active substances.

Furthermore, the focus of this guidance is on the technical mixtures of active substance(s) and their co-formulants undergoing an authorisation procedure with the Regulation (EC) 1107/2009. It is acknowledged that a number of formulations can be applied in close temporal proximity on the same crop or even together in the tank mix. This results in a huge number of possible combinations applied to the agricultural landscapes (each of them having different persistence in pollen and nectar), and boundless possible interactions between them. Bees are then typically exposed spatially and temporally to multiple residues (e.g., mix of insecticides, fungicides and herbicides). The guidance does not address the risk assessment of combinations of more than one PPP of unknown composition or when PPPs are applied sequentially within one season. However, EFSA has prepared a scientific opinion on a framework for a systems-based approach for the environmental risk assessment of PPPs for honey bee colonies (EFSA Scientific Committee et al., 2021). The system-based approach would allow in future to consider the multiple stressors and the complexity of the landscape of agricultural ecosystems that may impact on the bee health (More et al., 2021).

2.1 Structure of the Guidance Document

The guidance document includes several Chapters to address the various aspects of the risk assessment. Those are complemented by Appendices and Annexes. Some Appendices, i.e. Appendix A with the list of crop attractiveness and Appendix B with the parameters for contact and dietary exposure, are included as excel spread sheet.

The Annexes of the guidance document are:

Annex A	It gives guidance about refining residue dissipation, and has been developed in common with the guidance document of B&M;
Annex B	It includes recommendations for residue trials to refine the exposure estimation
Annex C	It includes recommendations for higher tier effect studies

In order to transparently document the science behind the revision of EFSA (2013), the guidance is complemented by a standalone Supplementary document, which includes all the background information, data collection and analysis. This Supplementary document includes different Annexes which were needed to report highly complex topics.

The Annexes of the Supplementary document are:

Annex A	Preliminary considerations and planned methods for the revision of Tier 1 risk assessment schemes of EFSA's 2013 guidance. This Annex provides an outline of the overall scientific processes that were considered to revise the risk assessments schemes in the new guidance by focusing on the problem formulation for Tier 1 risk assessment schemes, the related exposure scenarios (including crop attractiveness), the risk assessment parameters and the priority assigned to each of them. The implementation of this protocol is reflected in this document along with other topics that were considered for the new guidance
Annex B	Outcome of the systematic literature review on food consumption
Annex C	Outcome of the systematic literature review on the crop-specific sugar content in nectar
Annex D	Relevance of the flowering weeds scenario for the treated field
Annex E	Relevance of water scenario
ANNEX F	Expert Knowledge Elicitation (EKE) to assess the attractiveness of crops for pollen collection of bees
ANNEX G	Time-reinforced toxicity: concept and revised risk assessment scheme
ANNEX H	Residue database data (excel spread sheets)
ANNEX I	Succeeding crops
ANNEX J	Inter-species extrapolation data (excel spread sheets)
ANNEX K	Overview of sublethal effect testing on bees

3 Overview of the risk assessment

3.1 Specific Protection Goals

Within EFSA (2013), SPGs were defined according to the EFSA method mentioned in Section 1.2. For this guidance, a dialogue between risk assessors and risk managers was carried out through several consultations[1] to review the previous SPG definition, particularly in relation of the *magnitude* dimension, one of the 5 dimensions that define the overall level of protection according to the (EFSA Scientific Committee, 2016). The agreed SPGs that are implemented in this guidance document are reported in **Table 1**.

Table 1: overview of the agreed SPGs for honey bees, bumble bees, solitary bees

Dimensions	**Honey bees**	**Bumble bees**	**Solitary bees**
Ecological Entities	Colony	Colony	Population

Attribute	Colony strength	Colony strength	Population abundance
Magnitude	≤ 10%	Undefined	Undefined
Temporal scale	Any time	Undefined	Undefined
Spatial scale	Edge of field	Edge of field	Edge of field

529

530 The **ecological entity** dimension refers to the level of biological organisation for the
531 identified service providing units (SPUs), i.e. populations that deliver a given ecosystem
532 service (Luck et al., 2003). For honey bees and bumble bees, the ecological entity is '*colony*'
533 (i.e., a family of bees/a superorganisms consisting of one queen, several workers, and, in
534 some parts of the years, drones living in a single hive), while for solitary bees is '*population*'
535 *(*e.g. group of females and males of the same species nesting in the same place spatially
536 defined by the maximum foraging range). The ecological entity identifies the object of the risk
537 assessment both at lower tier and higher tier.

538 The **attribute** dimension allows identify the most ecologically relevant elements that must be
539 protected relative to the ecological entities. These elements are the main variables to be
540 addressed in the risk assessment and in experimental observations. For honey bees and
541 bumble bees the attribute is *colony strength*, which is defined operationally as the number of
542 adult bees it contains (= colony size). For solitary bees, the relevant attribute is the population
543 abundance.

544 The **magnitude** dimension refers to the level of tolerated effects for the attribute to be
545 measured relative to the defined ecological entities. This can be defined considering the
546 normal operating range (NOR) of the main variables identified for the attribute to be protected
547 and by considering the practical possibilities/limitation of measuring those variables in lower
548 and higher tier experiments. Following the dialogue between risk assessors and risk managers
549 and based on the scientific information provided by the EFSA WG (EFSA et al., 2021), risk
550 managers agreed on a magnitude dimension for honey bees (*A. mellifera*) for the entire EU
551 corresponding to a value of 10% as the maximum permitted level of colony size reduction
552 following pesticide exposure. For bumble bees and solitary bees, based on the consolidated
553 information provided in EFSA et al. (2022), an evidence-based decision for a threshold of
554 acceptable effects could not be finalised by risk managers due to the lack of data. The decision
555 was for an 'undefined threshold' that was given as an option in EFSA et al. (2022). However,
556 in order to ensure a better protection of these bee group, risk managers agreed that higher
557 tier data should be generated within approval of the active substances.

558 The **temporal scale** dimension defines the duration of tolerated effects. In principle, the
559 temporal scale for a threshold of acceptable effects is not relevant whenever a 'recovery
560 option' is not contemplated, since any possible effect following the exposure to a pesticide
561 should remain at a level indicated as acceptable at any time. However, it is noted that, in
562 practice, a temporal scale may be needed on the basis of practical limitations in the field
563 studies (i.e., the reference tier), since a continuous measurement of the main variables is not
564 practically feasible, nor is it advisable to inspect the colonies/nests too frequently, as this
565 creates stress for the bees which would affect the results of the experiments (see Annex C -
566 Recommendations for higher tier effect studies).

567 The **spatial** dimension 'edge of field' refers to the location of the colonies/populations, i.e.,
568 directly adjacent to the treated field, from where the bees forage in the treated field or

immediate off-field areas. Therefore, the different exposure scenarios considered in the risk assessment, refer to foraging in the treated field (e.g., treated crop, in-field flowering weeds, succeeding crop) and in the immediate areas surrounding the field (e.g., the field margin and the adjacent crop scenarios) (see Section 1.4).

This spatial scale definition is a worst-case exposure situation compared to colonies or populations located further away from the treated field. The WG also recognised that bumble bee and solitary bee nests located in the field may be exposed to pesticides, e.g. via direct exposure of larvae and adults to soil residues, contaminated nest materials. However insufficient information is available to address this exposure (Gradish et al., 2018, Sgolastra et al., 2019) and therefore to re-consider the definition of the spatial scale.

3.2 Implementation of the SPG in the risk assessment including tiered approach

3.2.1 General principles

The implementation of the agreed SPGs in the risk assessment requires the combined evaluation of the exposure generated by the use of a PPP in the field (which can be predicted, simulated, or measured) and of the ecotoxicological effects (which are assessed as part of the hazard characterisation based on an imposed exposure in the laboratory or higher tier effect experiments). To define in a structured and unambiguous manner, what exposure and which ecotoxicological effects should be used to implement the SPGs, the concepts of Exposure Assessment Goal (ExAG) and Effect Assessment Goal (EfAG) have been developed. The ExAGs relate to e.g., definition of the environmental exposure, type and duration (see Supplementary document for more details) and EfAGs relate to e.g., definition of relevant model species, type of toxicity endpoints.

The definition of the ExAG allows to answer questions such as

- where, in which matrix and for what time frame the exposure should be estimated, or
- what level of conservativeness the exposure estimate should aim for, i.e. what percentage of the exposure situations in the field should be covered in the risk assessment?

The definition of the EfAG allows to answer questions such as

- what should be the measured endpoints for the relevant species;
- what extrapolation approaches should be used to cover other species, endpoints and untested exposure regimes, or
- which percentile of a probabilistic effect assessment should be selected?

Bees will experience various levels of exposure at the edge of the treated field(s). This variability may be due to temporal differences (e.g. the same hive/nest may experience different exposure level in spring or during summer) or due to spatial differences (e.g. different hives/nests placed at different locations in the area of use of the PPP). Therefore, this means that it is necessary to define the ExAG, which can be determined, by selecting a percentile that will result in realistic worst-case exposure estimation from the distribution of the various levels of the exposures. Since a 90^{th} percentile is commonly used in ecotoxicology risk assessment e.g. for the EU FOCUS surface water, for this guidance document the 90^{th} percentile, already used in EFSA (2013) is retained.

It is noted that exposure and the effect (or hazard) assessments should address coherently the agreed SPGs in all the tiers and thus should be completely consistent with each other; for example, the effect assessment for honey bees focuses on hives located at the edge of treated fields, and thus the exposure assessment should not include hives located far from pesticide-treated agricultural areas.

The Ecotoxicologically Relevant Exposure Quantity[5] (EREQ) is the conceptual interface between the effect and exposure tiers. It is based on ecotoxicological considerations and defines the type of exposure quantity that in a mechanistic sense best explains observed effects in an ecotoxicological experiment. In general, the EREQ is defined as the residue intake per bee per time period, as given in an ecotoxicological experiment by the dose received per bee by dietary or contact exposure. On the exposure side, the EREQ can be quantified in form of the Predicted Exposure Quantity (PEQ), for example given for dietary exposure by pesticide intake per bee per time period determined from consumption and concentration in nectar and pollen in the field. Likewise, for contact exposure, the PEQ can be determined from respective exposure calculations. In the risk assessment, PEQ values are used then as input for effect calculations. Summarizing:

Terminology	Explanation
EREQ, Ecotoxicologically Relevant Exposure Quantity	Not a value, but a **type** of quantity, that gives the best mechanistic link between exposure and effects in an ecotoxicological experiment, and that is calculated/estimated both in the field (PEQ) and the ecotoxicological tests (dose)
PEQ, Predicted Exposure Quantity	A value; i.e. the quantification of an EREQ for a specific compound in the field/ area of use.
Dose	A value: administered exposure in ecotoxicological tests
EED, Estimated Exposure Dose	A value: estimated in effect field studies

3.2.2 Tiered approach

According to the ExAGs and EfAGs, both exposure estimation and effect assessment can be performed following a tiered approach, moving from relatively simple, conservative assessments to more realistic assessments. In fact, the concept of tiered approaches is to start with a simple assessment such as a screening assessment, or Tier 1 and add reality and complexity by moving to Tier 2, or higher tier, if necessary to refine the risk i.e. when the risk is not excluded at the lower tier. A fundamental aspect of the tiered approach is that every Tier of the exposure assessment should address the same ExAG, every Tier of the effect assessment should address the defined EfAG. Therefore, the same EREQ should be addressed in all tiers. For example, for the dietary exposure of bees all the exposure tiers should aim at estimating the 90th percentile of the daily intake averaged over all bees of the hive, and not a 90th percentile of the daily intake in lower tiers and a 10-d time weighted average value in a higher tier. For the risk assessment of bees exposure-tiers and effect-tiers are linked by

[5] In other documents also called ERC –Ecotoxicologically Relevant Concentration

using of the estimated exposure in the environment as the dose in the ecotoxicological test as is done in Chapter 7.

Both the ecotoxicological endpoints and the exposure in the field should be expressed as the same type of exposure quantity (e.g. µg/bee per day) in order to enable a consistent linking between each effect and exposure assessment tier.

Since both the exposure and effect assessments are operationalised in the tiered approach, it is appropriate to define an exposure-tier and an effect-tier:

- Exposure-Tier: In the exposure tiers, residue intake or residue deposition need to be quantified by calculating the PEQ to address the dietary and the contact route of exposure of the bees following the use of a PPP in the field.
- Effect-Tier: In the effect tiers the imposed exposure is called 'Dose' in the laboratory tests or 'Estimated Exposure Dose' in the higher tier tests.

An overview of the exposure-Tiers and effect-Tiers implemented in this guidance document is reported in Table 3.

For both exposure and effect assessment sides, the bee routes of exposure should be addressed by considering the different timescale of effects (acute and chronic) and the different life stages (adults and larvae). To this purpose in this guidance document four risk cases have been defined:

- Acute-contact
- Acute-dietary
- Chronic-dietary
- Larvae-dietary

The exposure estimation in the different Tiers will provide PEQ for each of the above risk cases and it is indicated as PEQ_j, where the suffix j indicates the four risk cases. In parallel, the effect assessment will provide $LD50_j$ and $slope_j$ which are used to connotate the dose-response relationship for each risk case (see Table 2).

Table 2: Risk cases defined in this guidance and related PEQ, LD50, slope

Risk cases	PEQ_j,	$LD50_j$ **and** $slope_j$
Acute-contact	$PEQ_{contact}$	$LD50_{contact}$, $slope_{contact}$
Acute-dietary	$PEQ_{dietary,acute}$	$LD50_{dietary,\ acute}$, $slope_{dietary,\ acute}$
Chronic-dietary	$PEQ_{dietary,chronic}$	$LDD50_{dietary,chronic}$, $slope_{dietary,chronic}$
Larvae-dietary	$PEQ_{dietary,larvae}$	$LD50_{dietary,chronic}$, $slope_{dietary,chronic}$

In the lower tiers of the exposure assessment, the exposure is based on default parameters, while in higher tiers the exposure of the colony (or population) may be based on measured parameters (e.g. PPP concentrations measured at the plant or brought into the hive/nest by bees) see Annex B - recommendations for residue trials to refine the exposure estimation).

Regarding the effect or hazard assessment, in the lower tiers it is based upon ecotoxicological experiments with individual bees in laboratory studies, while the highest tier is formed by different type of studies e.g. semi-field, colony feeder and/or field tests (see Annex C - Recommendations for higher tier effect studies). The field tests represent the highest level of

677 experimentally feasible effect assessments on bees foraging at local and larger scales (treated
678 field, immediate off-field areas and possibly landscape). In principle field tests can be
679 supported by population-level modelling of effects for different ecological and agricultural
680 practice scenarios, but such models first would need to be developed, calibrated and evaluated
681 for their use in regulatory bee risk assessment (see Section 10.7).

682 **Table 3:** tiered approach and explanations as to what each exposure or effect tier implies for
683 the risk assessment of active substance. According to the principle of the tiered approach,
684 each exposure-tier can be linked to each effect-tier

Exposure-Tier - Predicted Exposure Quantity (*PEQ$_i$*)					
Contact route of exposure	**Tier 1**			**Tier 2**	
	Based on default values of all parameters that take into account differentiation between exposure scenarios and body surface area of bees			Based on data to refine exposure values	
Dietary route of exposure	**Screening**	**Tier 1**		**Tier 2**	
	Based on default values of all parameters. No differentiation between scenarios, since a worst-case generic scenario is considered	Based on default values of all parameters that take into account differentiation between exposure scenarios, crops, time of application and residue behavior.		Based on different options for refinement of the default parameters.	
Effect-Tier					
	Tier 1		**Higher Tier**		
	All bees		**Honey bees**	**Bumble bees**	**Solitary bees**
Only Contact route of exposure	Based on standard toxicity endpoints and extrapolation factors for covering untested bee species. It includes geometric mean when more studies are available with the same species.		Semi-field studies	NA	NA
Only Dietary route of exposure	Based on standard toxicity endpoints and extrapolation factors for covering untested bee species. It includes geometric mean when more studies are available with the same species. For honey bees: • the chronic assessment requires considering the assessment of Time Reinforced Toxicity (TRT); • the acute and chronic assessment require considering the evaluation of potential concerns due to sublethal effects		Colony feeder studies **(larvae)**	Colony feeder studies **(larvae)**	NA
Combined contact and dietary routes of exposure			Field studies	Semi-field studies Field studies	Semi-field studies Field studies

	Population or colony models (see Chapter 10)

3.3 Risk assessment process

In this section the entire scientific process for the risk assessment included in this guidance document is described by presenting the various steps, indicating when additional data should be generated, when mitigation measure could be proposed and when it is possible to conclude. The overview for honey bees is given in the flowcharts in the Figure 1 (for the lower tiers) and Figure 2 (for the higher tiers).

An important and primary component of the risk assessment process is the definition of a proper problem formulation that, at lower tiers, would allow to identify the cases where a risk assessment is not needed as well as to frame the risk assessment where needed by selecting the more appropriate methodology (see Chapter 4) and, at higher tier, to select the more appropriate testing strategy (see Chapter 10).

Based on the problem formulation, when exposure to bees cannot be excluded, exposure estimation and effect assessment should be performed in order to identify the most relevant endpoints for risk assessment (see Chapters 5 and 6, respectively).

For the exposure, two quantities have been considered in this guidance, pending on the routes of exposure:

- Residue deposition (Red) for the estimation of the PEQ_j via contact exposure;
- Residue intake (Rint) for the estimation of the PEQ_j via dietary exposure.

Each of the above, corresponds to the PEQ_j for the exposure scenarios considered relevant following the problem formulation and the worst-case PEQ_j for each of the four risk cases needs to be selected. The same quantities are also used for estimating the PEQ_j at different exposure-tiers (see Table 3 above), in particular:

- Red will allow estimating the contact PEQ_j at Tier 1 and Tier 2.
- Rint will allow estimating dietary PEQ_j at screening, Tier 1, and Tier 2 .

As reported in 3.2.2 and in Table 3, the effect-tier assessment is based either on laboratory studies with individual bees (adults or larvae) at the lower tier or outdoor studies with colonies or populations (see Chapter 8) at the higher tier. The laboratory studies give the dose-responses for the different timescale of the effects (acute and chronic) and different life stages (adult and larvae) and therefore address the four risk cases mentioned above. They are the basis for identifying the endpoints ($LD50_j$ and $slope_j$) for honey bees, bumble bees and solitary bees. Higher tier studies, depending on the study type and on the organisms tested, provide a range of effect endpoints. In such studies, effects need to be investigated at an exposure level in line with the ExAG, i.e. 90th percentile worst-case exposure for the compound under evaluation. Therefore, appropriate exposure regimes and levels, defined as Estimated Exposure Dose (EED_j), have to be ensured in the study and expressed in the same unit as the PEQ_j (see Chapter 10). This means that, in addition to the biological observations, it may be necessary to verify consistently the exposure levels e.g. via measurement of residues in pollen and nectar, and use these measurements to estimate the EED_j in the specific study, which

need to be compared with the related PEQ_j in order to assess the plausibility of biological observations. The comparison should be carried out with PEQ_j based on independent measured residue trials (e.g. Tier 2), but if not available, the PEQ_j from lower tier exposure assessment will be used.

As part of the effect-tier assessment of the PPP, two additional aspects should in parallel be addressed at lower tier for honey bees: the potential for the compound under evaluation for showing increasing toxic effects due to long-term exposure to low doses, and sublethal effects i.e.:

- Time Reinforced Toxicity assessment (TRT) (see Chapter 8)
- Potential concerns due to sublethal effects (see Chapter 9).

TRT assessment is determined via extrapolation from the standard 10-day chronic honey bee toxicity study (OECD 245). It is important to note that when TRT is observed, it should be reflected by a proper selection of the honey bee chronic endpoint and by a proper risk assessment. In particular, both the toxicity endpoint and the exposure estimation (PEQ_j) should cover the whole lifespan of a honey bee and therefore, two scenarios for risk assessment were considered for covering the active period of the bees (i.e. 'summer bee scenario'), and an inactive winter period (i.e. 'winter bee scenario'). Regarding sublethal effects, the Tier 1 allows to identify potential concerns which should be addressed with further testing and/or go for a higher tier assessment with field studies, when those concerns cannot be excluded.

For lower tier risk assessment, the effects and the exposure estimations are reiteratively combined for the four risk cases to predict the expected risk at colony level (see Chapter 7).

It is recommended to initiate the lower tier risk assessment process with the screening or Tier 1 exposure assessments and Tier 1 effect endpoints and move to next exposure-tier only when appropriate, based on the outcome of the screening or Tier 1.

When an exposure-Tier 2 assessment is needed, applicants should generate appropriate data to replace the default values used in Tier 1.

When a low risk cannot be excluded based solely on the exposure refinement (i.e. exposure-Tier 2), a higher effect-tier assessment is required and applicants should generate studies investigating the effects under realistic worst case use conditions for the concerned bee group. In order to select the most appropriate testing strategy, the problem formulation should be reconsidered in light of the outcome of the lower tier risk assessment. For example, if, for the compound under evaluation, at the lower tier risk assessment a high risk is identified and the risk is driven by the effects on larvae, the higher tier risk assessment could be tailored to address this specific concern.

In this guidance, semi-field, field and colony feeder studies are suggested for honey bees and bumble bee, while semi-field and field studies are available for solitary bees (see Chapter 10). They represent different levels of realism and complexity and provide different endpoints that can be compared with the SPG for honey bees or evaluated analytically for bumble bees and solitary bees, since the magnitude dimension of their SPGs is not defined.

Any higher Tier risk assessment can, in principle, be supported by colony or population-level modelling. Such support could for example consist of using a honey bee colony model to simulate the exposure as observed in an exposure field study with the aim to understand or

extrapolate to other locations or other environmental conditions. In addition, effectiveness of suggested Risk Mitigation Measures (RMMs) for different ecological and agricultural practice scenarios could be assessed with the support of ecological population or colony models.

In any case, any model first would need to be evaluated and tested for use in regulatory bee risk assessment following good modelling practices (EFSA PPR Panel, 2014). In addition, the use of ecological modelling as higher Tier refinement method in isolation is not considered as an option.

EFSA outsourced the development of an agent-based model for a honey bee colony that was conceptualised by the MUST-B WG (EFSA, 2016). It is called **ApisRAM** and the first version of the formal model was recently published (Duan et al., 2022). Further implementation is currently under way for its possible use as a regulatory tool for the risk assessment of PPPs, but also for conducting more holistic risk assessments including exposure from multiple stressors at the landscape level (i.e. various infectious agents, pests, predators and other types of chemicals present in the hive or the environment, see ApisRAM timeline and data sources). More details on the possibilities for the use of ecological models are given in Chapter 10.

RMMs can be integrated to an exposure assessment re-estimation at any tier, except the screening level (See Chapter 13) and/or they can be proposed to reformulate the problem formulation. In both cases, RMMs should be proposed by the applicant and mentioned in the Good Agricultural Practices (GAP). When risk mitigation measures are integrated in the GAP, then the relevant exposure assessment should account for the suggested mitigation. In some cases, this will need the provision of exposure data whilst in other cases default values are available (e.g. spray-drift reduction) and can be used for the exposure assessment. What is essential is that any suggested mitigation is demonstrated to reduce the risk sufficiently so that the specific protection goal is met (i.e. a risk assessment indicating a low risk).

As part of the risk assessment process, also the risk from **metabolites** should be addressed. A risk assessment scheme is proposed on this guidance which has to be followed together with the evaluation of the PPP (see Chapter 11).

In this guidance, a proposal is also provided for the risk assessment of **mixture**. This is not routinely included as part of the risk assessment processes of the active substance, but it may be mainly relevant within the authorization of PPP at national level (see Chapter 12).

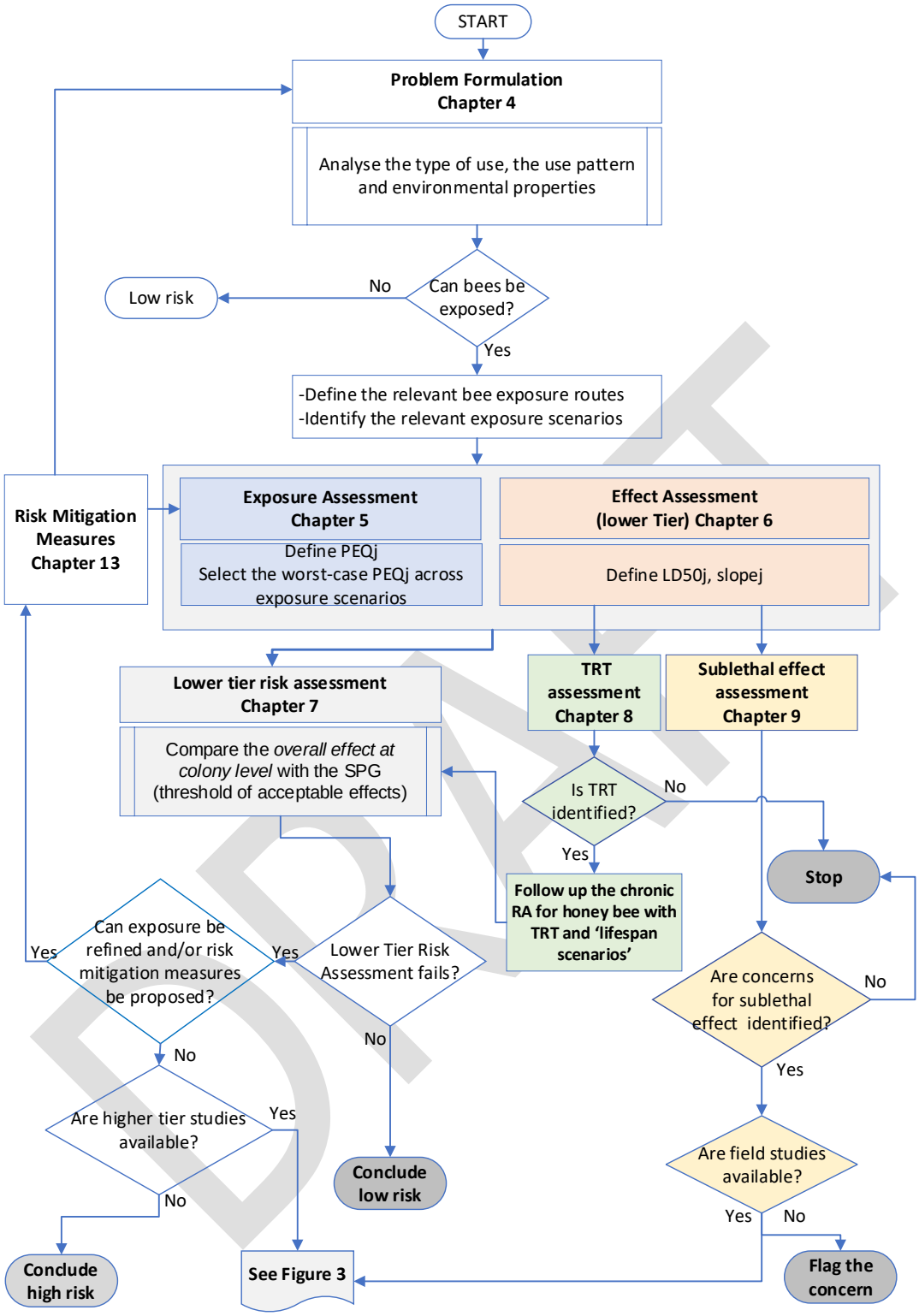


Figure 1: overview of the lower tier risk assessment process for the active substance for honey bees (HB). Risk assessment of metabolites covered in Chapter 11. Risk assessment of mixture covered in Chapter 12. TRT=time reinforced Toxicity (for honey bees). PEQj= Predicted Exposure Quantity for the four risk cases (indicated by the suffix j) i.e. acute-contact, acute-dietary, chronic-dietary, larvae-dietary.

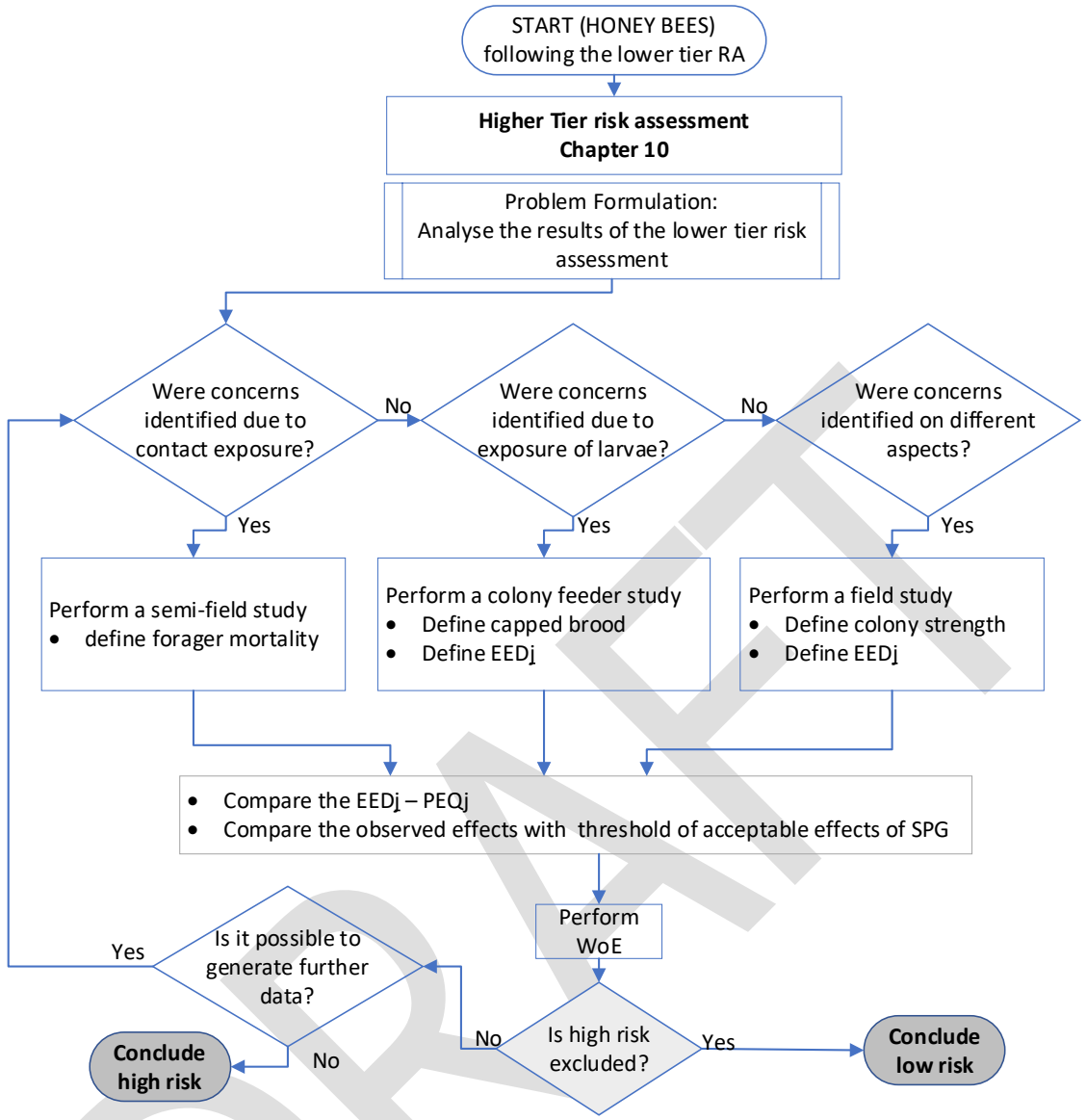


Figure 2: overview of the higher tier risk assessment scheme for honey bees (for further details see Chapter 10 and Annex C –)

4 Problem formulation

Problem formulation is the first step of the risk assessment which allows applicants and risk assessors to identify the potential hazard and exposure pathways for a PPP and to formulate risk hypotheses and identify the proper risk assessment methodology. The problem formulation sets the boundaries for risk assessment for making it 'fit for purpose'. For a risk to occur, it requires an exposure to a PPP which should result in a direct harm to the bees that exceeds a specified SPG. The use of PPPs may directly or indirectly affect bees, their habitat and or food availability. For example, a toxic effect on a bee would be considered a direct effect. Alternatively, the use of a herbicide may reduce the number of flowering weeds which in turn may reduce the available food in the bee's habitat, negatively impacting on individuals/colonies/populations of bees. The latter would be an example of indirect effect. This guidance document only considers direct effects.

When evaluating a substance/compound, it is important to first determine, through a **proper problem formulation**, if and how, based on its intended use, it could reach the bees and estimate the level of the exposure (see Section 5). Therefore, the starting point is a careful consideration of the proposed representative uses of a PPP as indicated in the GAP table and an overall consideration of the physical chemical properties and the behaviour of the compound in environmental matrices (mainly molecular weight, water solubility, partition coefficient octanol/water, dissociation coefficient, vapour pressure, soil sorption, persistence in soil, formation and transformation products). This includes an analysis of the methods of application, the crop(s) where the PPP is intended to be applied, the crop phenology (i.e. BBCH), the number of applications, the application rate and any particular conditions of PPP use (see Figure 3). Those aspects will allow applicants and risk assessors to understand the exposure pathways and, thus, to address the data requirements.

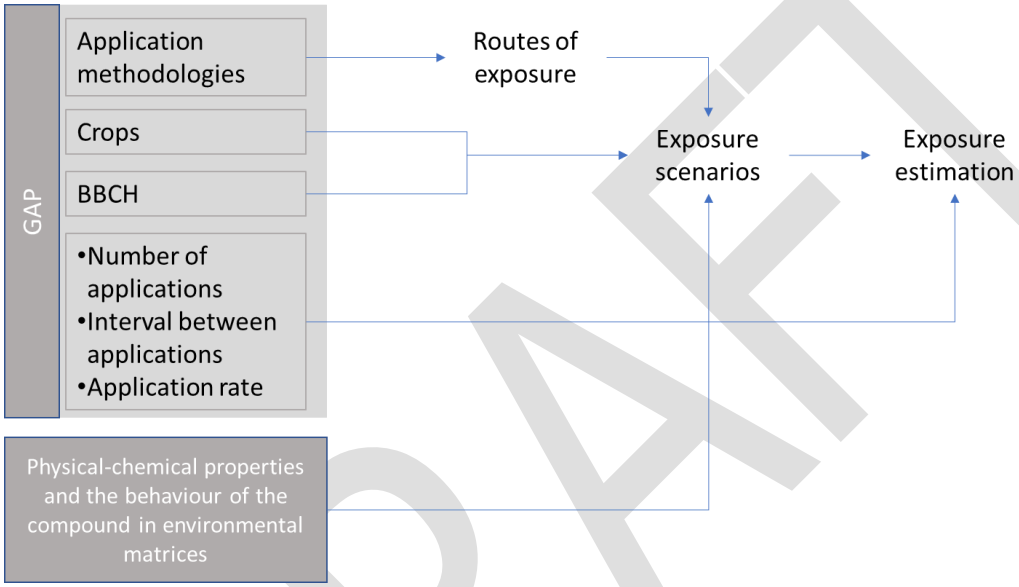


Figure 3: Relationships between the GAP/Phys-chem properties/behaviour of the compound in the environmental matrices and the bee exposure estimation to PPPs

Regarding the effects, if the exposure is not excluded, the problem formulation should consider which data are needed for the effect assessment relative to the relevant route of exposure including different type of the effects, timescales and lifestages (see Section 6).

4.1 Agricultural practices

A GAP table defines the way a PPP is proposed for use. It should include the following information (please note this list is not exhaustive):

- Information on the PPP including the product type (e.g., a water dispersible granule), which active substance(s) is included and at what concentration.
- The intended crops (or plants) and the growth stage of the crop (usually using the BBCH growth stage criteria). Sometimes the time of year when applications will be made may also be specified.
- The Member States or Regulatory Zones where it is intended that the PPP will be used.
- The intended application rate (normally in terms of amount of active substance per hectare but the rate of the product may also be stated).
- The number of applications and the application interval.

- Limited information on the method of application (i.e., broadcast air assisted sprayer, seed treatment etc.). The EPPO global database[6] for PPP treatments is a useful source of application methodology.
- Whether the PPP will be used indoors, in outdoor fields or in protected structures. According to EFSA (2014a) the type of protected structure should be defined (i.e., greenhouse or plastic tunnel).
- Risk mitigation measures proposed by the applicant.
- Other restrictions proposed by the applicant (e.g., applications are only allowed every three years).
- Sometimes other information on how the PPP is intended for use may be stated (e.g., band application or spot application).

There are numerous other agronomic- or growing-practices which will influence the presence, and exposure, of bees in areas where the PPP will be used. If details of such practices are specified in the GAP, then it allows a risk assessment to account for them and thus be more specific. Where information is lacking, a risk assessment should encompass (as far as practicable) all the conditions of where and how the PPP will be used. The following list includes the types of agronomic- or growing-practices which are expected to have an influence of the presence and/or exposure of bees (please note the list is not exhaustive):

- Diversity of crops in the landscape
- Orchard type (e.g., modern spindle, traditional)
- Presence and type of between-row vegetation.
- Crop development stage at time of harvest
- Time between the application and the start of the flowering for spray applications

4.2 Type of uses and application methodologies

During the problem formulation, it should be checked for each use in the GAP, whether the a.s./PPP under assessment is intended to be applied in open field or closed spaces.

The EFSA Guidance EFSA (2014a) and EPPO Global Database[6] use the following categories:

- Indoors and/or in permanent greenhouses (closed spaces);
- Semi-open structures (low mini tunnels, plastic shelters, net shelter/shade house and walk-in tunnels);
- Outdoor uses (open field).

For indoor uses and/or in permanent greenhouses, no exposure to bees is expected therefore a risk assessment is normally not necessary. However, it is noted that for applications made to seedlings, seeds or tubers which are subsequently transported to the field exposure to bees may occur. Equally for substances that are (semi-)volatile, exposure may occur following deposition of the substance in the proximity of the closed space. In these cases, a risk assessment may be required (See Table 4, note 8 and Chapter 5).

[6] https://gd.eppo.int/PPPUse/3CROLK

For semi-open structures and outdoor uses, exposure cannot be excluded and thus a subsequent consideration of the application methodologies is needed to understand how and where the bees can be exposed. Semi-open structures are assessed as outdoor uses.

The most relevant application methods that may lead to bee exposure are spraying of a liquid (emulsions, suspensions or solutions), seed treatments and distribution of granules. The present Guidance covers mainly these methods of applications, for which exposure estimation approaches are available and consolidated. However, bee exposure to PPP applied with other methods of application cannot be excluded. When a PPP is intended to be applied using an application method which is not covered by this Guidance including modern technologies, it is considered that the applicant has the responsibility to provide a proper characterisation of the exposure in line with the principles of this guidance (see Chapter 5).

An overview of methods of application covered by the guidance and other methods not covered is reported in Table 4 together with a consideration of the relevance of the routes of exposure (see Section 1.4) and exposure scenarios (see Section 4.3). The list was compiled based on the definition of EPPO nomenclature[7] and/or other existing uses registered in EU, although it may not be exhaustive.

[7] https://gd.eppo.int/taxon/3TMETM

903 **Table 4:** Overview of the possible type of uses and methods of PPP application in relation to the contact and dietary routes of exposure in the exposure
904 scenarios.

	Contact			Dietary				
	Treated crop	**Weeds* (treated field)**	**Field margin**	**Treated crop**	**Weeds (treated field)**	**Field margin**	**Adjacent crop**	**Succeeding crop / Permanent crop**
Seed treatment (coating)	Y	N	Y	Y	N	Y	Y	Y
Granular application	Y	Y	Y	Y	Y	Y	Y	Y
Conventional spray applications	Y	Y	Y	Y	Y	Y	Y	Y
Other spraying methods (i.e. aerial spray, in-furrow, knapsack, stem application, etc…) (See **note 1**)	Y	Y	Y	Y	Y	Y	Y	Y
Brushing *Application of a liquid product or powder with a brush*	N	N	N	Y	N	N	N	Y
Injecting *Application of a liquid product or solution by injecting it directly into the treated object*	N	N	N	Y	N	N	N	Y
Dipping *Application by immersing the treated object or part of it in a liquid product or solution* (See **note 2**)	N	N	N	Y	Y	N	N	Y

	Contact			Dietary				
	Treated crop	**Weeds* (treated field)**	**Field margin**	**Treated crop**	**Weeds (treated field)**	**Field margin**	**Adjacent crop**	**Succeeding crop / Permanent crop**
Drenching *Application of a liquid product or solution by pouring over the treated object* (See **note 3**)	N	N	N	Y	Y	Y	Y	Y
Dripping *Application of a liquid product or solution via multiple drops on substrate or soil*	N	N	N	Y	Y	N	N	Y
Placing *Application by positioning a product within target range* (See **note 4**)	N	N	N	Y	Y	Y	Y	Y
Circulating water application *Application of a product in the nutrient solution that it is circulated in a closed system to irrigate plants growing in substrates* *Presumed protected use*	N	N	N	N	N	N	N	N
Dusting *Application of a product by blowing tiny solid particles (dustable*	Y	Y	Y	Y	Y	Y	Y	Y

	Contact			Dietary				
	Treated crop	**Weeds* (treated field)**	**Field margin**	**Treated crop**	**Weeds (treated field)**	**Field margin**	**Adjacent crop**	**Succeeding crop / Permanent crop**
powder) towards the treated object								
Fogging *Application of a product by producing an atmosphere full of tiny droplets (particle size 0.05 - 50 microns)*	Y	Y	Y	Y	Y	Y	Y	Y
Fumigating *Application of a product that completely fills a confined space in a gaseous form* (See **note 5**)	Y	Y	Y	Y	Y	Y	Y	Y
Impregnating *Application of a liquid product or solution for absorption by a solid object* (See **note 6**)	N	N	N	N	N	N	N	N
Soil incorporation (as soil fumigants) (See **note 7**)	N	N	N	N	N	N	N	N
Permanent greenhouse (See note 8)	N	N	N	N	N	N	N	N

*The relevance of this scenario is also depending on the BBCH (see 4.3.2)

4.2.1 Notes to the possible type of uses and methods of PPP application

In this Section, additional remarks, mentioned as notes in Table 4 are reported.

As general remark, it is highlighted that for the uses reported in the Table 4 also particular methods of application such as band application, spot application, treatment between the row, etc. have been taken into account.

For uses in semi-open structures exposure cannot be excluded. Therefore, they must be considered as open field uses.

It is also pointed out that the information reported in Table 4 is not exhaustive, and it only considers the contact exposure via overspray or spray-dust-drift and dietary exposure routes via pollen and nectar. Pending on the method of application (i.e. soil fumigant, drenching, dripping, etc), bees may be exposed via other routes of exposure i.e. inhalation, and via other contaminated matrices, e.g. contact with contaminated soil or other materials used for nest building.

NOTE 1 on other spraying methods. The off-crop (i.e. field margin, adjacent crop) exposure of some of these methods of spraying may not be covered by the standard drift values used in the GD (e.g. for applications via helicopters, drones, etc..). Also, for other methods (e.g. stem application), flowers are not directly sprayed. For in-furrow applications, the product is applied together with the seed along the line drawn by the plough. Consequently, the exposure pathways differ for each of these methods and should be properly characterised.

NOTE 2 on Dipping. Bulbs, plants roots or entire seedlings are dipped in the product (or a solution of the product) before planting in the field. The seedlings have no flower at the time of application nor do they have them shortly after. A distinction should be made if soil surrounding the roots is also dipped into the product or just the bulb/roots. In such case, dietary exposure via the weeds may also be possible. Sometimes plant trays are dipped in the product.

NOTE 3 on Drenching. Drenching can be via a boom sprayer without the use of nozzles. In this case, there is no atomisation of the liquid and the majority of the liquid reaches the soil. No drift is assumed from this kind of application, but exposure might be possible via soil contamination. The outcome of any evaluation with this application technique is pending on the height and the accuracy of the device used for the application.

NOTE 4 on Placing. A solid object (rodlets, sticks, etc…) placed directly in the soil, beside the plants. Exposure might be possible via soil contamination.

NOTE 5 on Fumigating. Despite applied in a confined space, re-entry of workers may require opening the windows for ventilation. Off-crop exposure may then be relevant in some cases.

NOTE 6 on Impregnating. The applications of a liquid product for absorption by a solid object e.g. in nets impregnated with an insecticide or in traps in combination with an attractant are designed to kill pests. The potential of exposure should be properly characterized depending on the design (accessibility and/or attractivity).

NOTE 7 on soil incorporation (as soil fumigants). Soil fumigants may be injected (as liquids forming gas) into the soil and move through the soil mainly via diffusion in the gas phase. Exposure of bees in the air above and around the field of application is likely. Exposure to the off-crop via redeposition is likely if the treated soil is not properly sealed after application. This may lead to contamination of flowering plants growing outside the treated area (directly or via soil) and exposure of bees (contact and oral). A very specific exposure pathway is exposure via inhalation, which may be relevant for highly volatile substances such as soil fumigants under certain circumstances. Inhalation toxicity studies can then be an option for the risk assessment.

NOTE 8 on permanent greenhouse. Exposure cannot be excluded for items that are moved outside and for (semi-)volatile substances. If plants, seeds or tubers are moved outside after treatment, exposure via pollen and nectar may result from these treated items. For these uses, exposure to bees cannot be excluded and the risk assessment should be performed by considering the items moved outside for the treated crop scenario and succeeding crop scenario.

If the active substance is (semi-)volatile, deposition from the air to the area in the vicinity of the greenhouse or closed building may occur following venting. This may lead to contamination of flowering crops and plants growing outside the greenhouse/building (directly or via soil) and exposure of bees (contact and dietary). When deposition rate is calculated in the fate and behaviour section, this should be taken for the bee risk assessment.

4.3 Exposure scenarios

During the problem formulation, the most appropriate and relevant exposure scenario should be identified for the PPP and each use in the GAP for the bee exposure routes (e.g. contact and dietary exposure).

As explained in Section 1.4, bees may be exposed in the treated areas (i.e. treated crop and flowering weeds scenarios) and/or in the surrounding areas (i.e. field margin and adjacent crop scenarios). Furthermore, in some situations, bees may be exposed to residues in the pollen and nectar that are up-taken by the crops growing after the one under evaluation (i.e. succeeding crop scenario).

In relation to the contact exposure, in this guidance, it is considered that bees can be over-sprayed in the treated areas and/or could come in contact with spray drift or dust drift in the surrounding areas at the time of the PPP application.

In relation to the dietary exposure via consumption of contaminated pollen and nectar, it has to be noted that the **proportional contribution** of the various exposure scenarios to the daily food intake by bees is unknown. Therefore, the WG has retained the assumption of EFSA (2013), that each scenario contributes to 100% of the contaminated food consumed by bees, as worst-case. In theory, this might lead to stacking of selected exposure percentiles and thus extreme exposure probabilities. However, usually one of the exposure routes will strongly dominate the exposure, and thus, stacking of probabilities is unlikely to be an issue.

Pending on the GAP, during the problem formulation, the occurrence of exposure from some scenarios can be excluded a priori (see Table 4). In these cases, the exposure estimation is required only for the remaining relevant scenarios.

It is noted that among the most relevant scenarios, only those scenarios that will strongly dominate the exposure on the basis of the exposure estimation (see Section 5) will be used for risk assessment, since it is considered to cover all the others. This means that worst-case PEQ_j will be selected across scenarios for risk assessment (see Chapter 7). It is important to note that the 'worst-case PEQ_j should be identified at each tier of the risk assessment. For example, if for a substance/compound and its intended use under evaluation at Tier 1, the 'e.g. treated crop, or weed scenario' is identified as giving the worst-case PEQ_j ('dominant scenario') and the Tier 1 fails, at the Tier 2 all the scenarios should be reconsidered to redetermine the 'dominant scenario'. This will ensure that by including the refinement options the relevant 'dominant scenario' is always identified and the other scenarios are covered. Equally, when risk mitigation measures are applied at Tier 1 or Tier 2 (e.g. only apply outside the flowering period of the crop), the 'dominant scenario' should be redetermined.

4.3.1 Treated crop

The exposure of bees to PPPs requires that bees visit and interact with crops, therefore it is necessary to ascertain if crops in the GAP are attractive to bees. As pollen and nectar are the main sources of nutrition for bees, the WG defined a crop as being attractive based on the presence and availability of pollen and nectar. The WG decided that any crop that meets the following criteria is considered attractive to bees:

- The crop produces nectar which is accessible to bees
- The crop produces nectar and pollen which is accessible to bees

There is a third group of crops which produce pollen but no nectar. Within this third group, there are crops which are frequently visited by bees (e.g., tomato), while others are not. Following the publication of the previous Guidance Document (EFSA, 2013), there has been a debate on whether certain crops of this group should be considered attractive to bees. Therefore, to distinguish between these crops EFSA has performed an Expert Knowledge Elicitation (EKE) according to EFSA (2014b). (See Annex F to the Supplementary document for additional relevant information for the EKE project)

The details of the EKE are reported in Section 4.3.1 of the Supplementary document. Overall, 23 crops were identified for assessment in the EKE. Based on the results of the EKE, the WG has revised the list of attractive crops for bees and has made a new list that is available in the Appendix - A - Crop attractiveness.

When a crop is attractive to bees for pollen and/or nectar, contact exposure and dietary exposure cannot be excluded, pending on the BBCH stage of the crop. As reported in Table 5. When a crop is considered not attractive, the exposure is assumed to be zero and therefore the treated crop scenario is not relevant. Equally, the treated crop scenario is not relevant for the post-flowering treatment and for those crops which are harvested before flowering.

Table 5: Relevance of contact and dietary exposure for pollen/nectar attractive crops

	Attractive crop		
	Before flowering	Flowering	Post-flowering or harvested before flowering
Contact exposure	No	Yes	No
Dietary exposure	Yes	Yes	No

It is noted that for some crops, extrafloral nectaries may occur and consequently, exposure via nectar may be relevant outside the flowering period. Since no data are available to estimate this exposure, when relevant, it should be considered an uncertainty in the risk assessment.

If the GAP table refers to a crop which is absent from Appendix - A - Crop attractiveness, the applicant should propose a surrogate crop, for similar characteristics (morphology, phenology, etc..). The applicant should make it very clear that a surrogate has been used; they should also duly justify the choice of surrogate crop and how the selection was made. The selected surrogate crop will be risk assessors.

4.3.2 Weeds in the treated field

When the 'treated crop' scenario is considered not relevant for the bee exposure, bees may still be exposed in the treated areas while foraging on the flowering weeds present in those areas. The relevance of flowering weeds in the treated field as an exposure source for bees was assessed

based on the results from a re-analysis of the dataset from Last et al. (2019), in combination with the outcome of a consultation of efficacy experts. The full details on this exercise (i.e. dataset used, steps in the re-analysis, interpretation of the results, etc.) are reported in Section 4.3.2 of the Supplementary document.

Based on the available data and analysis, the WG concluded that for the dietary route of exposure, the flowering weeds scenario cannot be excluded *a priori* during the problem formulation. Pending on the GAP, the exposure estimation for this scenario is required to appreciate if it may drive the overall exposure for bees. For the contact exposure, the WG concluded that in some situations this scenario can be excluded a priori. When this is the case, there is no need to carry on with the exposure estimation. An overview of those cases where the flowering weeds scenario is or is not relevant for the contact route of exposure is provided in Table 6.

Table 6: Overall conclusion on the relevance of the flowering weeds scenario for the contact risk assessment for different crops and their respective BBCH stages at the time of application

Crop	BBCH stage (at the time of application)				
	0-9	**10-19**	**20-29**	**30-39**	**40-99**
Sunflower	No	No	NR	Yes [2]	Yes [2]
Maize	No	No	NR	Yes [2]	Yes [2]
Winter oilseed rape	No	No	Yes [2]	Yes	Yes
Winter cereals	No	No	No	Yes	Yes
Sugar beet [3]	No	No	NR	Yes [2]	Yes [2]
Potatoes	No	No	No	No	Yes [2]
Peas	No	No	NR	No	Yes [2]
Bean	No	No	No	No	Yes [2]
Other arable crops [1]	No	No	Yes [2]	Yes [2]	Yes [2]
Permanent crops	Yes				

[1] Including spring cereals and spring oilseed rape; [2] The data available is not sufficient for a conclusion, or no data is available at all. Therefore the flowering weeds scenario is considered relevant. This conclusion could potentially be revised when further data becomes available; [3]Unless cultivated for seed production, sugar beet is harvested at BBCH 49; NR: BBCH stage does not exist for this crop according to the BBCH Monograph (Meyer, 2001)

The WG acknowledged the limitations and uncertainties of the dataset and developed recommendations for further addressing these limitations (see Annex D of the Supplementary document). Although detailed guidance could not be provided for generating fit for purpose monitoring studies, some general indications for such studies were provided. The exposure estimation for this scenario is described in Chapter 5.

4.3.3 Field margin and adjacent crop

Areas surrounding the treated crop can be defined as field margin (wild vegetation) and adjacent crops (agricultural crops grown by a farmer in the neighbouring field). Pending on the GAP (see Table 4), the field margin as well as the adjacent crops have to be considered as a relevant exposure scenario for both the contact and the dietary route of exposure, since this represents a relevant area of interest for pollinator habitats. In these areas, deposition from PPP applications regularly occurs as spray-drift (spray applications), or dust-drift (after the sowing of treated seeds or application of granules). For these scenarios it is assumed that plants/crops are in flower when the drift/dust event occurs.

Taking into consideration biological, ecological and meteorological aspects, three set ups were recommended by the WG. These are the following (see also in Figure 4 and Table 7 below):

a. Field margin A: two metres widths and located immediately next to one of the sides of a rectangular treated field. It is assumed, that the field margin is always downwind. This

set up is to be used for the contact assessment for all the bees and for the dietary assessment for bumble bees and solitary bees.

b. Field margin B: two metres widths and located immediately next to all the four sides of a rectangular treated field. The consequence of this physical set up is that – irrespectively of the wind direction – one third of the field margin areas are contaminated due to spray drift or dust drift, while two thirds are located upwind thus remains uncontaminated. This set up is to be used only for the dietary assessment for honey bees.

c. Adjacent crop: it has 50 metres widths and is next to a rectangular treated field. It is assumed, that the adjacent crop is always downwind. This set up is to be used only for the dietary assessment for honey bees.

These considerations are reflected by the parametrization of the exposure assessment, that is described in Chapter 5.

It is noted that, as in EFSA (2013), no adjacent crop scenario for the contact exposure for honey bees is set as this scenario is covered by the field margin scenario. Further background and explanation on the three set up is included in 4.3.3 of the Supplementary Document.

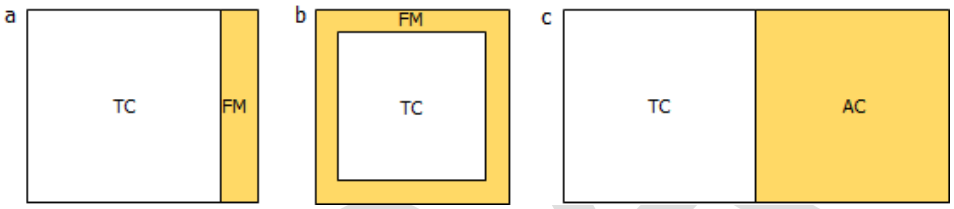


Figure 4: a - field margin A scenario for contact route of exposure for all bee groups and for the dietary exposure for bumble bees and solitary bees; b - field margin B scenario for the dietary exposure for the honey bees; c - adjacent crop scenario for the dietary exposure for honey bees (TC = treated crop; FM = field margin; AC = adjacent crop)

Table 7: summary of the relevance of the off-field scenarios

Scenario	Contact exposure	Dietary exposure
Field margin A	relevant for all bee groups	relevant only for bumble bees and solitary bees
Field margin B	not relevant	relevant only for honey bees
Adjacent crop	covered by field margin A, therefore not used	relevant only for honey bees, (covered by field margin A for the other bees)

4.3.4 Succeeding crop

In the succeeding crop scenario bees are exposed to pollen and nectar contaminated with residues of the substance (active ingredient and/or metabolites) that are already present in the soil following the treatment of the preceding crop. Residues that persist in soil are taken up by the roots of the succeeding annual crops or the permanent crops next year and then translocated via the vascular system and the tissues of plants to the nectar and pollen. This may also happen for the double-crops: annual crops that are grown twice in a growing season on the same field

(e.g. beans). Like in EFSA (2013), it is considered that if the succeeding crops are not defined, it is assumed that the crops are attractive for both the pollen and nectar.

The relevance of the succeeding crop scenario was re-examined based on the available field studies where the residues levels of a substance were measured in pollen and/or nectar collected from a crop grown as follow-on crop (see Section 4.3.4 of the Supplementary document). In addition, a screening level was established in order to identify those substances that, for a specific GAP, would lead to an exposure level that will not cause adverse effects on bees for the succeeding crop scenario. Based on the available data and analysis, the WG concluded that for the dietary route of exposure, the succeeding crop scenario cannot be excluded *a priori* but its relevance should be always considered during the problem formulation. However, for the annual double crops and for the permanent crops the following year, the WG concluded that in some situations this scenario can be excluded based on specific combinations of soil persistence and soil adsorption properties of a substance with toxicity endpoints ≥ 0.1 µg/bee and at a given application rate (Table 8 and Table 9). It should be noted that a necessary condition for the applicability of the screening level is that all the toxicity endpoints (i.e. all the acute LD50 values, the LDD50 and the larval ED50) must be ≥ 0.1 µg/bee, ≥ 0.1 µg/bee/day and ≥ 0.1 µg/larva/developmental period. When this is the case, then there is no need to further assess the succeeding crop scenario.

Table 8: Screening level for the relevance of the succeeding crop exposure scenario based on different combinations of soil persistence and adsorption properties of a substance and application rates (expressed as total annual application) to permanent crops. The screening level is applicable only when all the toxicity endpoints (i.e. all the acute LD50 values, the LDD50 and the larval ED50) are ≥ 0.1 µg/bee.

Application rate ≤ 100 g/ha	**Application rate ≤ 500 g/ha**	**Application rate ≤ 1 kg/ha**	**Application rate ≤ 5 kg/ha**	**Application rate ≤ 10 kg/ha**
DT50 ≤ 3 days Koc ≥ 100 mL/g	DT50 ≤ 3 days Koc ≥ 100 mL/g	DT50 ≤ 3 days Koc ≥ 100 mL/g	DT50 ≤ 3 days Koc ≥ 100 mL/g	DT50 ≤ 3 days Koc ≥ 100 mL/g
DT50 ≤ 10 days Koc ≥ 500 mL/g	DT50 ≤ 5 days Koc ≥ 500 mL/g	DT50 ≤ 5 days Koc ≥ 500 mL/g		
DT50 ≤ 30 days Koc ≥ 2000 mL/g	DT50 ≤ 10 days Koc ≥ 2000 mL/g	DT50 ≤ 10 days Koc ≥ 5000 mL/g		
DT50 ≤ 60 days Koc ≥ 5000 mL/g				

Table 9: Screening level for the relevance of the succeeding crop exposure scenario based on different combinations of soil persistence and adsorption properties of a substance and application rates (expressed as total annual application) to annual "double" crops. The screening level is applicable only when all the toxicity endpoints (i.e. all the acute LD50 values, the LDD50 and the larval ED50) are ≥ 0.1 µg/bee.

Application rate ≤ 100 g/ha	**Application rate ≤ 500 g/ha**	**Application rate ≤ 1 kg/ha**	**Application rate ≤ 5 kg/ha**	**Application rate ≤ 10 kg/ha**
DT50 ≤ 3 days Koc ≥ 100 mL/g	DT50 ≤ 3 days Koc ≥ 100 mL/g	DT50 ≤ 3 days Koc ≥ 500 mL/g	DT50 ≤ 2 days Koc ≥ 100 mL/g	DT50 ≤ 2 days Koc ≥ 100 mL/g
DT50 ≤ 5 days Koc ≥ 500 mL/g	DT50 ≤ 5 days Koc ≥ 2000 mL/g	DT50 ≤ 5 days Koc ≥ 5000 mL/g	DT50 ≤ 3 days Koc ≥ 2000 mL/g	DT50 ≤ 3 days Koc ≥ 5000 mL/g

DT50 ≤ 10 days Koc ≥ 2000 mL/g				
DT50 ≤ 30 days Koc ≥ 5000 mL/g				

Chapter 5 provides guidance for the exposure assessment for the three types of succeeding crop scenarios.

5 Exposure Assessment

5.1 The exposure assessment models

The exposure assessment models (i.e. the mathematical expression of the exposure estimation) described in this Chapter are to be used to predict the exposure quantity of an individual bee for the two main routes of exposure; the contact route of exposure and the dietary exposure arising from the application of a PPP. These predictions are to be used in the lower tier risk assessments (see Chapter 7). The model to be used for the contact route of exposure is included in Section 5.1.1 and the models to be used for the dietary exposure are in 5.1.2. In 5.1.2, two models are included. The scope and the applicability of those models are explained in the respective sections.

Other routes of exposure could be relevant in some situations such as inhalation (see Chapter 4). However, exposure models and default parameters are not available. Therefore, applicants should generate fit for purpose data for characterising both exposure and hazard.

5.1.1 Contact model

Contact exposure is a form of acute exposure that requires physical contact between the PPP and the surface of the bee (see Section 1.4). This may happen during or shortly after the PPP application, therefore this route of exposure is relevant for foraging honey bees, foraging worker bumble bees, and adult solitary bees.

The lower tier exposure assessment from EFSA (2013) has been reviewed and updated as it is presented below:

$$RED = PEQ_{co} = AR\ Ef_{co}\ Bsf \qquad \text{[eq.1]}$$

Where:

RED: Residue deposition
PEQ_{co}: Predicted Exposure Quantity for contact exposure - µg/bee
In other parts of the Guidance document, PEQ is indexed with a 'j' since PEQ is used for different risk cases ('j' is referring in general to the risk cases, here represent the acute contact risk case and is indicated by the suffix 'co')
AR: application rate – g/ha
Ef_{co}: exposure factor for contact exposure (-)
Bsf: body surface factor - dm^2/bee

The detailed descriptions of the parameters are included in Section 5.2.

The contact model as described above is applicable for all types of application methodologies covered in this Guidance Document, for all relevant scenarios and for all bee groups.

5.1.2 Dietary models

The dietary exposure considers that bees or the bee larvae come to contact with the PPP via consumption the contaminated food (pollen and or nectar). As indicated in Section 1.4, acute and chronic dietary exposure is considered in this Guidance Document for the adult bees and chronic exposure for the larvae.

The exposure assessment considered in the lower tiers by EFSA (2013) has been reviewed by the WG and several amendments were proposed. In addition to the reparameterization of the existing parameters, the way of using some of the parameters has been reconsidered and additional parameters have been introduced. That resulted in two models considering numerous parameters which, however, are able to estimate the quantity of the PPP residue intake by an individual in a more accurate way. The first model predicts the residue intake when the contamination of pollen and nectar dominantly originates from contamination of the above-soil parts of the crop/plant (see Section 5.1.2.1). The second model predicts the residue intake when the contamination of pollen and nectar originates from the soil (see Section 5.1.2.2).

In this Guidance Document, for PPPs that exhibit time reinforced toxicity (see Chapters 6 and 8), some special scenarios that are called as 'summer bee' and 'winter bee' are considered. The exposure estimations for those scenarios are detailed in the respective sections.

5.1.2.1 *Dietary model for above-soil contamination*

The dietary model for the above soil contamination is to be used (for all risk cases) when the crop or the plant might directly be contaminated by the PPP. This might happen in case of the following scenarios and situations:

- treated crop scenario when the PPP application is made after the emergence of the crop (BBCH ≥ 10)
- weeds in the field scenario when the PPP application is made when the weeds are already emerged
- field margin scenario
- adjacent crop scenario

This model has been set up in the following way:

$$Rint = PEQ_{di} = AR\ Ef_{di}\ (PFF\ (SV_{po,be} + SV_{ne,be}) + SV_{po,du} + SV_{ne,du}) \quad \text{[eq.2]}$$

The SV parameters are derived by using the following equations:

$$SV_{po,be} = \frac{LF_{po}\ CONC_{po,be}\ CMP_{po}}{1000} \quad \text{[eq.3]}$$

$$SV_{ne,be} = \frac{LF_{ne}\ CONC_{ne,be}\frac{CMP_{su}}{SN}}{1000} \quad \text{[eq.4]}$$

$$SV_{po,du} = \frac{LF_{po}\ CONC_{po,du}\ CMP_{po}}{1000} \quad \text{[eq.5]}$$

$$SV_{ne,du} = \frac{LF_{ne}\ CONC_{ne,du}\frac{CMP_{su}}{SN}}{1000} \quad \text{[eq.6]}$$

The CONC parameters are derived by using the following equations:

$$CONC_{po,be} = f(RUD_{po}, n_{be}, i_{be}, DT50_{pnt}) \quad \text{[eq.7]}$$

$$CONC_{ne,be} = f(RUD_{ne}, n_{be}, i_{be}, DT50_{pnt}) \quad \text{[eq.8]}$$

$$CONC_{po,du} = f(RUD_{po}, n_{du}, i_{du}, DT50_{po}, w) \quad \text{[eq.9]}$$

1202 $$CONC_{ne,du} = f(RUD_{ne}, n_{du}, i_{du}, DT50_{ne}, w)$$ [eq.10]

1203 Short definitions for each parameter are summarized in a Table 10, further below, and the
1204 detailed descriptions with the parametrizations are available in Section 5.3.

1205 5.1.2.2 *Dietary model for through soil contamination*
1206 Dietary model for through soil contamination is to be used (for all risk cases) when the crop or
1207 the plant that consist the scenario can only be contaminated via the soil. This might happen in
1208 case of the following scenarios and situations:

1209 • treated crop scenario when the PPP application is made before the emergence of the crop
1210 (BBCH < 10)

1211 • weeds in the field scenario when the PPP application is made when the weeds are not yet
1212 emerged, i.e. the GAP is explicitly for 'bare soil' situation

1213 • succeeding crop scenario

1214 This model has been set up in the following way:

1215 $$Rint = PEQ_{di} = SV_{po,\mathrm{soil}} + SV_{ne,soil}$$ [eq.11]

1216 The SV parameters are derived by using the following equations:

1217 $$SV_{po,\mathrm{soil}} = \frac{LF_{po}\, PEC_{pw}\, CMP_{po}}{1000}$$ [eq.12]

1218 $$SV_{ne,soil} = \frac{LF_{ne} PEC_{pw} \frac{CMP_{su}}{SN}}{1000}$$ [eq.13]

1219 Short definitions for each parameter are summarized in **Table 10**, and the detailed descriptions
1220 with the parametrizations are available in Section 5.3.

1221 **Table 10:** Parameters of the dietary exposure model for above-soil contamination

Parameter	Definition	Unit*
Rint	Residue intake by an individual bee or bee larva	ng/bee or ng/bee/day or ng/larva/developmental period
PEQ_{di}	Predicted Exposure Quantity due to dietary exposure, i.e. the intake of pesticide mass per bee. This is the output of the exposure estimation. For the chronic adult assessments, this quantity has to be expressed per day, but for the larvae it has to be expressed as the sum of the intake over the entire developmental period. In other parts of the Guidance document, PEQ is indexed with a 'j'. 'j' is referring to the risk case like acute, adult chronic or larvae.	
AR	Application rate	g/ha
Ef_{di}	Exposure factor for dietary exposure	- (unitless)
PFF	Pre-Flowering Factor	- (unitless)
$SV_{po,be}$	Shortcut value for pollen for before flowering situations (product of eq. 3)	µg/bee or µg/bee/day or µg/larva/developmental period
$SV_{ne,be}$	shortcut value for nectar for before flowering situations (product of eq. 4)	µg/bee or µg/bee/day or µg/larva/developmental period
$SV_{po,du}$	shortcut value for pollen for during flowering situations (product of eq. 5)	µg/bee or µg/bee/day or µg/larva/developmental period
$SV_{ne,du}$	shortcut value for nectar for during flowering situations (product of eq. 6)	µg/bee or µg/bee/day or µg/larva/developmental period
$SV_{po,soil}$	Shortcut value for pollen for situations for contamination from soil (product of eq. 12)	µg/bee or µg/bee/day or µg/larva/developmental period

$SV_{ne,soil}$	Shortcut value for nectar for situations for contamination from soil (product of eq. 13)	µg/bee or µg/bee/day or µg/larva/developmental period
CMP_{po}	pollen consumption	mg/bee or mg/bee/day or mg/larva/developmental period
CMP_{su}	sugar consumption	mg/bee or mg/bee/day or mg/larva/developmental period
SN	sugar content of the nectar	kg/kg (i.e. -)
LF_{po}	landscape factor for pollen	- (unitless)
LF_{ne}	landscape factor for nectar	- (unitless)
$CONC_{po,be}$	concentration in pollen from before flowering application (product of eq. 7)	mg/kg
$CONC_{ne,be}$	concentration in nectar from before flowering application (product of eq. 8)	mg/kg
$CONC_{po,du}$	concentration in pollen from during flowering application (product of eq. 9)	mg/kg
$CONC_{ne,du}$	concentration in nectar from during flowering application (product of eq. 10)	mg/kg
RUD_{po}	residue unit dose of pollen	mg/kg
RUD_{ne}	residue unit dose of nectar	mg/kg
$DT50_{po}$	half-life in pollen; the time within which the concentration in pollen is reduced by 50 percent	day
$DT50_{ne}$	half-life in nectar; time within which the concentration in nectar is reduced by 50 percent	day
$DT50_{pnt}$	half-life in plant matrixes; the time within which the concentration in crop/plant matrixes is reduced by 50 percent	day
n_{be}	number of applications before flowering	- (unitless)
n_{du}	number of applications during flowering	- (unitless)
i_{be}	interval between multiple applications performed before flowering	day
i_{du}	interval between multiple applications performed during flowering	day
w	time window for deriving time-weighted average concentrations for chronic exposure	day
PEC_{pw}	Predicted Environmental Concentration in pore water	mg/L

*for some parameters, three possible units are given. When the exposure for the acute timescale is to be estimated, the pollen and sugar consumption in mg/bee has to be used. This will result in respective shortcut values in µg/bee. For the chronic adult assessment, the unit for the food consumption is mg/bee/day resulting in shortcut values in µg/bee/day. In the assessment for the larvae, the unit for the food consumption is mg/larva/developmental period and the related shortcut values will be in µg/larva/developmental period.

It should be noted that the exposure estimation in EFSA (2013) was expressed in µg/bee, but here it is in ng/bee. This is related to the way of considering the application rate, which was expressed in kg/ha in EFSA (2013), but it is now in g/ha. This change has three consequences: 1) the unit to be used for the application rate is now harmonized with the unit used for the contact module (both are now in g/ha); 2) the resulting PEQ_{di} values will be in a more comfortable order when compared to the methods used in EFSA (2013); 3) special care has to be made when the different PEQ_{di} values, obtained from lower tier or higher tier exposure estimations, are compared since some will be in ng while others might be expressed as µg. Also, special care has to be made when PEQ_{di} values are considered together with PEQ_{co} value, as the latter is in µg/bee.

5.2 Contact exposure parameters – Tier 1

5.2.1 Application rate (AR)

The application rate is the mass of the PPP applied to a certain size of the surface area of the treated field and it is the basis of the exposure-Tier estimations in the various scenarios. The unit of the application rate used in this guidance is gram per hectare (g a.s./ha).

The application rate to be considered is as it is reported in the GAP table for the PPP under the assessment. Since the higher the application rate is the higher the resulting estimated exposure, when the application rate is expressed as a range, the risk assessment should be conducted by considering the highest possible application rate. Using the highest possible application rate would cover the entire GAP. Contrary, using a lower value from the possible range would cover only partially the GAP. If the application rate is reported in other unit(s) than mass per area (e.g. g a.i. per kg seed for a seed treatment), they have to be converted into g/ha. Therefore, applicants are strongly encouraged to clearly report the application rate expressed in mass per area, or provide all necessary information that allows the risk assessor to express the application rate in such a unit.

5.2.2 Exposure factor for the contact exposure (Ef_{co})

Although the exposure estimation is dependent on the application rate in all cases, the exposure level of the bees is pending on the source of the exposure from the landscape (see 1.4). The Guidance Document considers different scenarios (see 4.3) to reflect this. The role of contact exposure factor (Ef_{co}) parameter is to quantify the differences in the exposure of the different scenarios. The parameters for Ef_{co} are derived from deposition factors. The deposition factor for the weed scenario is related to the crop interception (i.e. dependent on the growth stage of the crop) and the deposition to the field margin is related to the spray drift/dust drift.

All the details regarding how the values for the Ef_{co} are derived are included in Section 5.2.2 of the Supplementary Document and the Ef_{co} factors to be used in the contact risk assessment are reported in Appendix B –.

5.2.3 Body surface factor (Bsf)

This parameter was not considered in EFSA (2013), but is introduced here.

This parameter was introduced in order to get individual level exposure estimations that considers the size differences between the bee species to be covered in the exposure assessment. Therefore, a set of factors that is related to the surface of the bees was established. All the details regarding the methods used and the performed calculations are included in Section 5.2.3 of the Supplementary Document. The values to be used for the risk assessments are summarized in Table 11, below.

Table 11: The body surface factors to be used in the exposure assessment

Category for the risk assessment	Representative species	Bsf (dm^2/bee)
Honey bee	*Apis mellifera*	0.0114
Bumble bee	5th percentile (by body surface) bumble bee species	0.0146
Solitary bee	5th percentile (by body surface) solitary bee species	0.00184

5.3 Dietary exposure parameters – Tier 1

5.3.1 Application rate (AR)

The considerations for the application rate to be used in the dietary model are the same as for the contact model as described in 5.2.1, above.

5.3.2 Exposure factor for the dietary exposure (Ef_{di})

Similarly, to the considerations in 5.2.2, the role of the dietary exposure factor (Ef_{di}) is to quantify the differences in the exposure of the different scenarios.

The parameters for Ef_{di} are derived from three factors: deposition factor, dust formation factor and safety factor. The latter is a correction factor to extrapolate form spray-drift to dust-drift reported in EFSA (2013). The relevance of one or the other factors and the belonging to parameter is depending on the application method and the scenario. The deposition factor for the weed scenario is related to the crop interception i.e. dependent on the growth stage of the crop) and the deposition to the off-field scenarios, it is related to the spray drift/dust drift.

The details of the derivation of this parameter, are included in Section 5.3.2 of the Supplementary Document and the Ef_{di} factors to be used in the dietary risk assessment are reported in Appendix B.

5.3.3 Pre-flowering factor (PFF)

This parameter was not considered in EFSA (2013), but is introduced here.

The Pre-Flowering Factor (PFF) parameter was introduced with the revision of the GD (see Annex A of the of the Supplementary Document). This parameter has the function extrapolating from the estimated pollen and nectar concentrations from spray applications made in BBCH 60-69 (i.e., during flowering) to concentrations from spray applications made in BBCH 10-59. In order to establish this factor, several processes were considered. All details are included in Section 5.3.3 of the Supplementary Document. The default values to be used for the Tier 1 exposure estimations are summarized in Table 12, below.

As a summary, the proposed PFF considers only two underlaying mechanisms, the crop interception and the dissipation of the pesticide in plant tissues (in addition a default value is considered). In the first step the applied PPP is considered to follow two pathways: the soil deposition and the plant deposition (see Section 1.4). In the second step, the part which deposits to the soil contaminates the pollen and nectar via plant uptake. This is accounted for by considering the worst-case default RUD of 1 mg/kg (see more in Section 5.3.15). The part which deposits on the plants dissipates with the default DT_{50} of 10 days (see Section 5.3.10) and the remaining residues contribute in the contamination of the pollen and the nectar. Finally, the residues from the two routes are summed up and compared with the 90^{th} percentile RUDs at BBCH 60-69 (see Section 5.3.8); the ratio of the two results in the unitless PFF factor. Since the process in the plant is time dependent, PFF is defined in relation to the number of days that elapse between the spray application and the start of the flowering. Therefore, in order to be able to use this factor in the risk assessment, this number of days has to be defined. Since the length of this period might depend on several aspects and might vary considerably across different EU regions, the WG recommends to apply a worst-case approach in a first instance. The shorter the period is, the more conservative the resulting exposure estimation will be. In case the estimation for the number of days between the spray application and the start of the flowering is not available, the PFF category that assumes a period of less than 15 days has to be considered.

In this category, the exposure assessment has to be performed with the same parameters as used for spray applications during the flowering.

Table 12: The pre-flowering factors (PFF) to be used in the risk assessment for spray applications performed before the flowering

Time period between the last application and the start of the flowering				
> 50 days (PFF category 5)	49-35 days (PFF category 4)	34-25 days (PFF category 3)	24-15 days (PFF category 2)	< 15 days (PFF category 1)
0.08	0.09	0.17	0.33	no PFF, conduct the exposure estimation if it was for BBCH 60-69

It is noted that PFF is applicable only for spray applications made in BBCH 10-59 period and is applicable only for the treated crop scenario.

5.3.4 Shortcut values ($SV_{po,be}$, $SV_{ne,be}$, $SV_{po,du}$, $SV_{ne,du}$)

A shortcut value (SV) is the 90th percentile of the distribution of the residue intakes, based on an application rate of 1 kg/ha. The distribution is defined across the spatial statistical population of all colonies (or populations) at the edge of treated fields.

The methodology of the derivation has been extensively reviewed (see Section 5.3.4 in the Supplementary Document). Several amendments were made to the underlaying equations (Section 5.1.2 and eq. 3-6) and additional parameters were included (see Annex B – of this document).

The ranges of variability of the parameters were also revised. Based on the new equations and parameter ranges, the SVs were re-calculated with a Monte Carlo method (similarly to EFSA (2013)). The values were calculated separately for nectar and pollen, before and during flowering (see eq. 3-6 for $SV_{po,be}$, $SV_{ne,be}$, $SV_{po,du}$, $SV_{ne,du}$ in Section 5.1.2.1).

Note that the Rint (Eq 2) is calculated combining the four SV, i.e. the four 90th percentiles, which represents a conservative estimation of the true spatial 90th percentile of the Rint.

5.3.5 Food consumption (CMP_{po}, CMP_{ne})

Food consumption during the active period for bees

The active period of the yearly cycles of the bees that is characterized by intensive foraging activity, reproduction and brood care coincides with the vegetation period and most of the pesticide uses in open-field cropping systems. For that period, a systematic review and an extensive review were conducted to screen the available data from the scientific literature on the sugar/carbohydrates and pollen consumption rates by larvae and adults bees (honey bees, bumble bees and solitary bees). Considering the collected data, the WG has derived the food consumption values to be considered for the risk assessments. All details of the data collection, the results and the summary of the WG discussions are included in Section 5.3.5. of the Supplementary Document. The values to be used for the risk assessments are summarized in Table 13 and Table 14, below. Where a range of values are reported, a uniform distribution was considered for the shortcut value calculations (see Section 5.3.4).

Table 13: Food consumption of adult bees during the active period for bee

Category for the risk assessment	Representative species and bee role category	Daily sugar consumption (mg/bee/day)	Daily pollen consumption (mg/bee/day)
Honey bee	*Apis mellifera* forager	acute: 42 – 83 chronic: 0 – 83	0
	A. mellifera nurse	34	11.6
Bumble bee	5th percentile (by body weight) generic model bumble bee species	acute: 42 – 84 chronic: 0 - 84	11.7
Solitary bee	5th percentile (by body weight) generic model solitary bee species	acute: 2.2 – 4.5 chronic: 0 – 4.5	0.6

Table 14: Food consumption of bee larva during the active period for bees

Category for the RA	Representative species	Sugar consumption over the development period (mg/larva/developmental period)	Pollen consumption over the development period (mg/larva/developmental period)
Honey bee	*A. mellifera* larva	81.5	1.52 - 2.04
Bumble bee	*Bombus terrestris* larva	194.6	60.23
Solitary bee	*Osmia bicornis* larva	91	80.7 - 92.5
	O. cornuta larva	165	80.7 - 92.5

It has to be noted that the presence of open flowers is a necessary condition for the pollen and nectar consumption, i.e. consumption happens in BBCH 60-69 period of the crop or the plant. It is considered that after this period, the pollen and nectar production is stopped and the food consumption from that crop/plant drops down to 0 mg.

Food consumption during the inactive period for bees

Bumble bees and solitary bees do not store food for the inactive period; the overwintering forms are dormant in that period without food consumption. However, honey bees are known to consume carbohydrates even in the inactive winter period. No data collection was conducted for winter honey bees during the review process. However, relevant information was available in the EFSA opinion (EFSA PPR Panel, 2012). EFSA PPR Panel (2012)considered the sugar consumption of 8.8 mg sugar/bee/day (connected to thermoregulation) with no pollen consumption for the inactive winter period. The WG agreed to use this information where needed (e.g. in Chapter 8).

5.3.6 Sugar content of the nectar (SN)

The sugar content of the nectar is crop/plant-dependent, but varies also due to a number of abiotic factors. Since the sugar content of the nectar determines the energetic value of the nectar, the lower sugar content results in higher nectar consumption. Therefore, the sugar content significantly influences the exposure via nectar consumption. A systematic review was conducted to review the available data from the scientific literature on the sugar content of crops grown in the EU. Considering the collected data, the WG has allocated the crops into four sugar content categories and has defined the sugar content values to be considered for the risk assessments. All details of the data collection and the results are included in Section 5.3.6 of the Supplementary

Document and its related Annex (Annex C). The values to be used for the risk assessments are summarized in Table 15, below. Crops not present in this table belong to the category with the lowest sugar content, i.e., category 1; resulting in values of 10% for solitary bees and 15% for honey bees and bumble bees to be considered in Tier 1 exposure assessment. The default 10% and 15% values are the same as set in EFSA, 2013a (based on some relevant information on nectar quality foraged by bees). In case of the adjacent crop scenario and the succeeding crop scenario, the crop type is not defined. Therefore, in the Tier 1 risk assessment, the default worst case values of 10% and 15 % has to be used (i.e. category 1). As in EFSA (2013), weeds in the field and the field margin scenarios are considered as habitats with mixed vegetation and the sugar content of 30% is to be considered for the risk assessment.

Table 15: The allocation of the EU crops into sugar content categories and the sugar content values (SN) to be used in Tier 1 risk assessments for the different bee groups

Crop group	**Sugar category**	**Sugar content in nectar (%)**		
		Honey bees	**Bumble bees**	**Solitary bees**
Anise, badian, fennel, corian, apricots, pears, chillies and peppers, lemons and limes, tobacco	1	15	15	10
Almonds, blueberries, buckwheat, cherries, chicory roots, currants, grapefruit (inc. pomelos), leguminous for silage, melon, mustard seed, oranges, plums and sloes, pumpkins, squash and gourds, quinces, rapeseed, raspberries (and similar berries), safflower seed, seed cotton, sour cherries, soybeans, strawberries, turnips for fodder, vetches	2	20	20	20
Alfalfa, apples, beans, broad beans, horse beans (dry), clover for forage and silage, cucumbers and gherkins, peaches and nectarines, peppermint, phacelia, sesame seed, sunflower seed	3	30	30	30
Cabbages and other brassicas, onions	4	40	40	40

5.3.7 Landscape factor (LF_{po}, LF_{ne})

This parameter was not considered in EFSA (2013), but is introduced here.

Since bees are mobile species, they can visit and forage a large number of patches and cropped fields on the landscape. This factor describes the proportion of the food intake of a bee colony or population that originates from the treated field. A narrative data collection was undertaken by the WG. Sufficient and appropriate data for the food collection from the landscape could be compiled only for honey bees collecting pollen. All details about the methods, the data analysis, the results are included in Section 5.3.7 of the Supplementary Document. As a summary, the pollen share from 24 fields and landscapes (honey bee hives located at the edge of the field with

flowering attractive crop) had been grouped. Some information about the trials with the maximum proportion of pollen collected from the fields are reported in Table 16.

Table 16: Maximum crop pollen collected per field (highest crop pollen percentage for one hive at one sampling point of all the hives and timepoints at that field)

Crop	number of hives on a field	number of samplings in time	Maximum proportion of pollen originating from the crop (%)	Value to be considered for the SV calculations (-)
oilseed rape	8	2	70	0.70
oilseed rape	8	2	70	0.70
oilseed rape	8	2	69	0.69
oilseed rape	8	2	72	0.72
oilseed rape	4	3	58	0.58
oilseed rape	4	3	95	0.95
oilseed rape	4	3	59	0.59
oilseed rape	4	3	66	0.66
oilseed rape	4	3	53	0.53
oilseed rape	4	3	60	0.60
oilseed rape	4	3	84	0.84
oilseed rape	4	3	81	0.81
oilseed rape	4	3	69	0.69
oilseed rape	6	3	100	0.10
oilseed rape	6	3	100	0.10
oilseed rape	6	3	100	0.10
oilseed rape	6	3	100	0.10
oilseed rape	6	3	83	0.83
oilseed rape	6	3	85	0.85
oilseed rape	48[a]	2	98	0.98
oilseed rape	48[a]	2	98	0.98
Phacelia tanacetifolia	4	3	76	0.76
Phacelia tanacetifolia	4	3	75	0.75
Phacelia tanacetifolia	4	3	66	0.66

a: There were six fields close together in each test region, with 8 hives per field. As the fields were not independent, the regions were considered the relevant unit in this study.

Since the dataset included cases with up to 100% crop pollen collection of single sampling date, the WG considered that the landscape factor for the acute Tier 1 exposure assessments should be 1 (100% of the collected pollen origin from the contaminated area). The WG decided that all figures reported in Table 16, above (as a range) are to be considered in the Tier 1 shortcut value (SV) derivation only for the SVs calculated for pollen and for the honey bee adult chronic and the honey bee larva. For all other cases LF of 1 will be considered in the Tier 1 exposure assessment.

5.3.8 Residue Unit Dose (RUD)

The residue unit dose (RUD) is the parameter expressing the residue concentration of the pesticide molecule in pollen and in nectar, standardised on an application rate of 1 kg/ha. RUD

values to be used for the risk assessments – namely used for the Tier 1 shortcut value calculations (see Section 5.3.4) were derived from supervised residue trials. As reported in Annex A of the Supplementary Document, most of those trials had been collected and a database was built up before the review process has started. The database was published in 2017 as an external scientific report (Kyriakopoulou et al., 2017). In the frame of this review, the existing database was amended with initial residue values (and residue decline estimations, see in Section 5.3.9, below) from dossier studies submitted to EU regulatory bodies under Regulation (EC) 1107/2009 in the period between 2017-2019. As already indicated in Annex A of the Supplementary Document, the amended database includes sufficient data only for spray applications. All details of the amendments, the data analysis and the results are included in Section 5.3.8 of the Supplementary Document. The default values to be used for the shortcut value calculations are summarized in Table 17, below.

Table 17: RUD distributions to be considered for the shortcut value calculations

Matrix	Direction	Median (mg/kg)	SD (log scale)	Correlation between multiple applications (log scale)
Pollen	Downward (DW)	67.7	1.15	0.83
Nectar	Downward (DW)	0.87	2.06	0.84
Pollen	Sideward/upward (SUW)	192.6	0.71	0.39
Nectar	Sideward/upward (SUW)	10.3	0.30	0.85

It has to be noted that RUD values are the concentrations in mg/kg for an application rate of 1 kg/ha. There is no specific data set for pollen and nectar concentrations after dust drift contamination. Therefore, as in EFSA (2013), the default values (as presented above) for downward (DW) spraying are to be considered for granular and seed dressing applications for estimating the concentrations in pollen and nectar after dust drift contamination (for those situations a safety factor is applied as described in Section 5.3.2). In addition, it has to be noted that the available residue data (therefore the default RUDs) are relevant to be considered only for those mechanisms (source of exposure) when the chemical is deposited onto the pollen and nectar from the air (i.e. direct contamination of the pollen and the nectar; this mechanism is relevant only in the flowering stage of the crop/plant). The other way of pollen and nectar contamination is through the plant matrixes (indirect contamination). This has two main pathways: 1) the chemical deposits to plant surfaces other than the flower and from those surfaces it infiltrates and distributes in the plant tissues; 2) the chemical deposits to the soil, then the chemical is taken up by the roots of the plant from the soil and then the chemical distributes in the plant tissues. These two mechanisms are taken into consideration in the PFF parameter as described in Section 5.3.3. The RUDs as described above in combination with the PFF is used in situations when the PPP is applied between BBCH 10 and 59. All these above is illustrated in Table 18 below.

Table 18: Overview of the use of the different default RUDs for the different situations

Tier 1 scenario	spray applications	granular application	seed treatment
Treated crop, BBCH ≥ 10	default RUDs for DW or SUW spray application pending on the type of application	default RUDs for DW spray application	no RUDs are used*

Weeds in the field	default RUDs for DW spray application, but for bare soil situation (weed BBCH < 10) no RUDs are used*	not relevant
Field margin	default RUDs for DW spray application note: default assumption is that the plants at the field margin are in flowering stage	
Adjacent crop	default RUDs for DW spray application note: default assumption is that the crop is in flowering stage	

* no RUDs are considered as for those situations the 'trough soil dietary model' is considered (see in 5.1.2.2)

5.3.9 Half-life in pollen and nectar ($DT50_{po}$, $DT50_{ne}$)

These parameters were not considered as standalone parameters in EFSA (2013), but they were part of the TWA (time weighted average) parameter.

The database that had been used for residue unit dose derivation (see 5.3.8) included a number of trials that allowed an estimation for the residue decline in pollen and nectar. Therefore, this analysis was undertaken. All details about the methods, the data analysis, the results and the summary of the WG discussions are included in Section 5.3.9 of the Supplementary Document. As a result, the agreed default values to be used in Tier 1 risk assessments are the following:

- half-life in pollen ($DT50_{po}$): 3 days
- half-life in nectar ($DT50_{ne}$): 2 days

As in EFSA (2013), these values are relevant only for spray applications and are to be used only for certain chronic assessments for which a certain time window was established (see 5.3.13, below). This is because no residue decline information was available after dust drift contamination of pollen and nectar or for indirect contamination (through plant matrixes). Residue decline is not considered for acute assessments.

5.3.10 Half-life in plant matrixes ($DT50_{pnt}$)

This parameter was not considered in EFSA (2013), but is introduced here.

The EFSA Guidance Document on risk assessment for birds and mammals (EFSA, 2009) recommends a default value of 10 days for residue decline on sprayed plant foliage. Although the 10-day value for plant DT50 is routinely used in the risk assessment for birds and mammals, it may be argued that this value is not conservative enough to be implemented at Tier 1 in the risk assessment for bees. The WG decided to further investigate the adequacy of the default plant DT_{50} value of 10 days for the use in this Guidance Document. For that purpose, a narrative review was undertaken. The details of this review including the methods, the data that had been considered and the outcome of the exercise are included in Section 5.3.10 of the Supplementary Document.

In summary, relevant data and databases were identified that included a huge number of decline data in and on plant matrixes. The available data was considerable also in terms of number of compounds and plant species, variety of plant components and tissues and diversity of environmental conditions. In conclusion, the WG was of the opinion that the use of the default DT_{50} value of 10 days is sufficiently protective to be considered for the Tier 1 exposure assessment for bees.

5.3.11 Number of applications (n_{be}, n_{du})

These parameters were not considered in EFSA (2013).

The number of applications before flowering (n_{be}) and the number of applications during flowering (n_{du}) are simply the number of applications according to the GAP of the PPP for which the risk assessment is conducted. If the number of applications is reported as a range, it is advised to start the risk assessment with the highest number of applications since this will cover situations with lower number of applications (i.e. the entire GAP will be covered). In case of high risk conclusion for the risk assessment with the maximum number of applications, the risk assessment may be conducted with lower number of applications in order to conclude specifically for those situations. However, those risk assessments will not cover the entire GAP.

Considerations for of the treated crop scenario

Ideally, the GAP table should include clear information about the distribution of the applications between the pre-flowering (BBCH < 60), during flowering (BBCH 60 - 69) and after flowering (BBCH > 70) periods. If this is not the case, then, as a worst-case assumption, the risk assessment might be conducted by assuming that all the applications are performed during the flowering period. In certain situations (high number of applications combined with long intervals between individual applications), that assumption might be considered as unrealistic or it reflects an unfeasible application regime. In those situations, further information/considerations on the distribution of the applications has to be collected. The WG considered that, as a general rule, the length of the flowering period can be maximized in 28 days. Therefore, the number of applications during the flowering can be maximized to the number which fits to this 28-day period considering the length of the interval between applications (e.g. only 2 applications with 21-day interval will fit within the flowering window even if the total number of applications in the GAP is higher)In case of uncertainty, the WG recommends to apply the general worst-case approach. For that, it should be considered that applications conducted during the flowering period – in general -result in a higher exposure than the exposure predicted for the before flowering situations. Moreover, no exposure is expected for the after flowering applications. Appropriate information on the length of the flowering period of the crop, might also be considered.

For seed treatment, the number of applications is by default 1 and it is at BBCH 00.

Considerations for of the weed in the field scenario, the field margin scenario, adjacent crop and succeeding crop scenario

As default, it is assumed that the crops/plants that consist of those scenarios are in flowering stage at the time of the PPP application. Therefore, the WG recommended to consider that maximum number of applications as indicated in the GAP should be allocated to the parameter n_{du}. Consequently, the parameter $n_{be,}$ will be by default 0. It is acknowledged, that this approach might be worst case in situations with a high number of applications for the crop scenarios. Therefore, it is recommended to report clear and details information in the GAP.

5.3.12 Interval between multiple applications ($i_{be,}$ i_{du})

These parameters were not considered as standalone parameters in EFSA (2013), but the interval between multiple applications performed during the flowering was part of the TWA (time weighted average) parameter. The interval between multiple applications performed before flowering is a new parameter introduced here.

These parameters are not relevant for GAPs with single application per season, but for multiple applications before or during the flowering period of the crop or multiple applications in both time periods. The interval between multiple applications performed before flowering (i_{be}) and the interval between multiple applications performed during the flowering (i_{du}) are simply the number

of days that elapse between two applications according to the GAP of the PPP under evaluation. If the interval is reported as a range, it is advised to start the risk assessment with the lowest number of days since this will cover (worst-case) situations with longer intervals. In case of high risk conclusion for the risk assessment with the minimum number of days between two applications, the risk assessment may be conducted with longer possible interval(s) in order to conclude specifically for those situations. However, those risk assessments will not cover the entire GAP.

Ideally, the GAP table should include clear information about the interval(s) between the applications at the different periods. In case of uncertainty, the WG recommends to apply the general worst-case approach. For that, it should be considered that the shorter interval result in a higher exposure. For the weed in the field scenario, for the field margin scenario, for the adjacent crop and for the succeeding crop scenario always the shortest interval has to be considered irrespectively whether it belong to i_{be} or i_{du} (it is expected however, that in most of the cases i_{be} and i_{du} will be equal).

5.3.13 Time window (w)

The time window to be used in the risk assessment is based on the same logic as the time weighted average used in the previous guidance document (EFSA, 2013). This parameter is the number of days over which time-weighted average concentrations are considered for the chronic exposure. The parametrization proposed in this guidance is based also on the previous guidance (EFSA, 2013). As in EFSA (2013), in most of the cases, the proposed time window is one day (default). This is the case for all acute assessments, for all pre-flowering assessments and also for all assessments for solid formulations (seed treatment, granules). The parameters to be used for the spray applications performed during the flowering are presented in Table 19, below.

Table 19: The time window parameter (w) to be used in the risk assessment for spray applications performed during the flowering

Category for the risk assessment	Time scale of the effects and exposed life stage	Time window (day)
Honey bee	chronic adult	10
	chronic larva	5
Bumble bee	chronic adult	10
	chronic larva	1
Solitary bee	chronic adult	10
	chronic larva	1

The WG has considered whether for the chronic larva assessments for bumble bees and solitary bees the time window parameter could be set differently from the default, which was considered as too severe. However, in lack of relevant biological/ecological knowledge on all the species (such as time needed for the preparation of the provision, length of larval development) or the residue behaviour in the provision, no alternative proposals could be made.

5.3.14 Shortcut values ($SV_{po,soil}$, $SV_{ne,soil}$)

As described in 5.3.4, a shortcut value (SV) is the 90th percentile of a distribution of residue intake per bee (or larvaa) over a colony (or population, for solitary bees). For the calculations of $SV_{po,soil}$, $SV_{ne,soil}$ values, the same method as described in 5.3.4 are used. The only difference is

that in this case, the equations 12 and 13 are used as reported in Sections 5.1.2.2, above. (See Appendix B – of this document).

5.3.15 Predicted concentrations in soil pore water (PEC_{pw})

PEC_{pw} parameter for the Tier 1 exposure estimation is a single default value, which is 1 mg/kg.

However, it has to be noted that this Tier 1 exposure estimation based on the default value of 1 mg/kg, is applicable only for GAPs where the cumulative application rate is not higher than 4.5 kg/ha. More details regarding the limit of 4.5 kg/ha are included in 5.3.15 of the Supplementary Document.

In case, where the cumulative application rate is higher than 4.5 kg/ha, Tier 2 exposure estimation have to be conducted that requires a GAP specific PECpw calculation, as described in 5.5.15.

5.4 Contact exposure refinement – Tier 2

In the subsections of this Chapter it is explained which Tier 1 parameters can be refined for a Tier 2 exposure estimation and methods for the refinement of those parameters are recommended. Certain parameters cannot be refined. In those cases the same parameter value as defined for Tier 1 has to be used for the Tier 2 estimations. Tier 1 and Tier 2 exposure estimations use the same model as described 5.1.1.

It should be noted that when exposure studies for contact exposure are available (see Annex B), the values obtained from those experiments will represent the PEQ_{co} values themselves, representing a GAP situation equivalent for the use of the PPP in the experiment.

5.4.1 Application rate (AR)

The application rate is defined by the GAP of the PPP for which the risk assessment is performed. Therefore, no refinement option is possible.

Nevertheless, as indicated for Tier 1, if the application rate is expressed as a range and the risk assessment with the highest recommended application rate indicates a high risk, the risk assessment might be conducted by considering a lower rate in range. However, it has to be noted that this risk assessment would cover only partially the GAP.

5.4.2 Exposure factor for the contact exposure (Ef_{co})

Since crop interception values are available only for a limited number of crops/crop categories, many crops had to be grouped with the existing categories in order to set the Tier 1 values for Ef_{co} for the weeds in the field scenario. The crop interception is, however, crop specific as it depends on the morphology and growing pattern of the crops. Differences in growth intensity between species within the group might exist. In addition, novel varieties, novel cultivation or application technics could also result in some inaccuracy in the Tier 1 Ef_{co} values. Due to the high variability of potential influencing factors, no generic guidance can be provided for the refinement of the Ef_{co} value. Nevertheless, applicant can propose another Ef_{co} value for a Tier 2 risk assessment. This is the case also for crops that might not be covered by this Guidance (i.e. no Tier 1 Ef_{co} values proposed). The proposal should, ideally, be supported by experimental data or information from pertinent literature, but in any case, has to be duly reported. In those situations, the risk assessor has to carefully consider and decide, which deposition category or deposition value is the most appropriate for their particular case.

A number of crops were linked to two spray drift categories and the higher value was set to be the Tier 1 Ef_{co} value for the off-field scenarios. In cases when this approach is considered to be too conservative (e.g. the PPP use is restricted to a specific growing structures or spaying

technics), another Ef_{co} value for a Tier 2 risk assessment might be proposed. Again, this refinement will be case specific, therefore no generic guidance can be provided.

Special attention is needed for ornamentals, since ornamental plants are a diverse group of plants, grown in a variety of ways, which can vary from small herbaceous plants to large ornamentals trees. For this reason, the selection of appropriate parameters for environmental risk assessment is not straightforward. Ornamentals do not have their own deposition categories (neither for crop interception nor for spray or dust drift). The applicants and risk assessors should consider what is the most appropriate surrogate crop/crop category to use for the risk assessment taking into consideration the plant structure, plant density as well as the growth stage, the application methodology and other relevant information. The above instructions may be applied for crops grown in nursery as well as for some other crops.

5.4.3 Body surface factor (Bsf)

The body surface factors are dependent on the bee species only. Therefore, no refinement option is possible.

5.5 Dietary exposure refinement – Tier 2

In the subsections of this Chapter, it is explained which Tier 1 parameters can be refined for a Tier 2 exposure estimation and methods for the refinement of those parameters are recommended. Certain parameters cannot be refined. In those cases, the same parameter value as defined for Tier 1 has to be used for the Tier 2 estimations. Tier 1 and Tier 2 exposure estimations use the same model as described in 5.1.2.

It is important to note that refinement of the RUD parameters derived from field measurements with pre-flowering applications, cannot be applied together with refinement of the PFF. Similarly, – pending on the sampling method used for the RUD refinement – the RUD refinement cannot be applied together with the refinement of LF. In those cases, if refined RUD values are used in a Tier 2 assessment then both the PFF and the LF must take the parameter value of 1. This is because all the three parameters affect the residues levels entering the hive/nest, and refinement of those parameters in combination could lead to double counting processes. Nevertheless, if the sampling for the RUD refinement was performed in a way that the landscape effect was excluded, LF refinement would still be possible. All this above illustrated in the matrix below:

	RUD refinement performed by samples taken from plant or from bees/pollen traps in semi-field conditions*	RUD refinement performed by samples taken from bees/pollen traps in field conditions*
Combination with PFF refinement possible?	No (the effects of the underlaying process of PFF are reflected in the RUD refinement if it was for pre-flowering application)	
Combination with LF refinement possible?	Yes (the sampling method prevents landscape effects)	No (the sampling method includes landscape effects, albeit limited in case of sites with minimum alternative forage)

*further explanations/definitions are available in Section 5.5.8.

Refined PFF and LF parameters can be combined in a Tier 2 estimation with the Tier 1 RUD parameter values.

5.5.1 Application rate (AR)

The same considerations as described in 5.4.1, above are applicable here, as well.

5.5.2 Exposure factor for the dietary exposure (Ef_{di})

The same refinement options as proposed for Ef_{co} are applicable for Ef_{di}. Those options are described in 5.4.2 above.

5.5.3 Pre-flowering factor PFF

The residue levels in pollen and nectar from pre-flowering spray applications largely depend on the fate and behaviour of the PPP in the environment, including its fate and behaviour in/on the plants. Many of the underlaying process (e.g. root uptake, uptake via the leaves, mobility in plant, accumulation in plant matrixes, excretion to pollen/nectar, any elimination process, etc.) could be quantified by appropriate methods and used for refinement of the generic Tier 1 parameter for PFF. However, it would go beyond the remit of this Guidance Document to present detailed methodology for all those options. Nevertheless, the WG recommends further research to study or even develop models for the residue behaviour of the PPPs in plants, including prediction for residue concentration in pollen and nectar after the PPP application.

The WG, however, recommends three options which might be used to calculate a refined PFF parameter for Tier 2 exposure estimations.

- reconsider the crop interception assumption
- refine the dissipation rate in plant matrixes
- refine the contribution of pollen and nectar residue levels from the soil route

It should be noted, that if RUD refinements for pre-flowering applications are used for a Tier 2 exposure estimation (see in Section 5.5.8), then no refinement of the PFF is acceptable (and it must take the parameter value of 1 in the Tier 2 residue intake estimation).

Reconsider the interception assumption

As explained in 5.3.3, the Tier 1 PFF considers the crop interception (CI) in order to split the applied pesticide to soil route and plant route. The crop interception for the Tier 1 PFF were set in a generic, non-crop specific and non-BBCH specific way, resulting in some conservative assumptions. Namely, most of the Tier 1 PFF categories considers that 90% of the applied pesticide follows the plant route (CI = 90%). These Tier 1 PFF categories are the category 2, 3 and 4 (in these situations the contribution of the soil route is rather small). In the 5th category, the importance of the soil route is considerable, therefore here 10% crop interception is taken into consideration. In a Tier 2 exposure estimation, crop and BBCH specific crop interception value might be considered. For the existing crop interception values, Section 5.3.2 might be consulted.

Refine the dissipation rate in plant matrixes

The Tier 1 PFF takes into consideration the dissipation of the active substance in plant matrixes. In Tier 1, a default worst-case parameter value is considered (see in Sections 5.3.3 and 5.3.10). However, the consideration of active substance and crop specific decline data could lead to more realistic exposure estimation. Guidance on the derivation of active substance and crop specific decline data is provided in Section 5.5.10. For the calculation of a Tier 2 PFF, the methods for the Tier 1 calculations (for the plant route) as described in Section 5.3.3.2 of the Supplementary Document should be followed.

Refine the contribution of pollen and nectar residue levels from the soil route

The contribution of the soil route in the pollen and nectar contamination is considered by using a default worst-case parameter value in the Tier 1 PFF (see in Section 5.3.3). The contamination from this route is proportional with the crop interception. As described in Section 5.5.15, the pollen and nectar residue levels are considered to be proportional to the porewater concentration in the soil at the time of flowering. For a Tier 2 exposure estimation, the GAP specific porewater concentration can be modelled by available calculator tools as described in Section 5.5.15. It is important to note that the modelled porewater concentration considers the crop interception (PERSAM Tier 2). If GAP specific crop interception (see above) and the calculated pore water concentration are considered together in a Tier 2 exposure estimation, than the CI values has to be aligned.

5.5.4 Shortcut values ($SV_{po,be}$, $SV_{ne,be}$, $SV_{po,du}$, $SV_{ne,du}$)

The following parameters contributing to the shortcut values might be refined: LF, SN, RUD, DT50, w (note: 'w' could be refined, however not recommended; for details see 5.5.13). In case of appropriate Tier 2 data are available for one or more of those parameters (see corresponding sections of this Chapter), the SV should be recalculated by considering the Tier 2 data. The other, not refined parameters, must take the Tier 1 parameter value. The resulting values are considered to be Tier 2 shortcut values. The method of calculation is the same as described in 5.3.4.

5.5.5 Food consumption (CMP_{po}, CMP_{ne})

The WG has identified significant knowledge gaps regarding the food consumption of bees and bee larvae (see Section 5.3.5). Therefore, for some of the Tier 1 food consumption values bear some uncertainties. Even so, the WG do not recommend refinement of this parameter for a specific risk assessment. This is because the food consumption rates are general parameters belonging to the bee species (bee categories used for the risk assessment) and they are considered as independent from the PPP itself or from the use of the PPP (i.e. the GAP).

5.5.6 Sugar content of the nectar (SN)

EFSA (2013) recommends that data on crop specific sugar content can be considered as a refinement for Tier 2 risk assessments. Appendix S of EFSA (2013) prescribes a study design where at least 5 varieties of the same crop, each at least in 5 fields, should be sampled for sugar content determination. However, these recommendations from EFSA (2013) were only supported by a narrative literature review considering a limited number of pertinent studies. Therefore, the WG undertook a review of the recommendations given in the previous Guidance Document (EFSA, 2013). The details of this data analysis are included in 5.5.6 of the Supplementary Document. The analysis showed that the range of sugar content measured in an experiment conducted according to the requirements set in EFSA, 2013a would have a considerable overlap with the true hypothetical range in most of the cases. Therefore the WG has agreed that the recommendations as described in EFSA (2013) (Appendix S of (EFSA, 2013) are to be kept as recommendation in the reviewed Guidance Document. In summary, the WG considered that a refinement of the SN parameter for the treated crop scenario is recommended if suitable and sufficient field data are available. This is for any crop, in any Tier 1 sugar content category. As described in EFSA (2013), at least five varieties have to be involved and the sugar content of the nectar in at least five locations for each variety has to be available. For further definition for 'location' Chapter 10, has to be consulted. The WG recommends direct sampling from the nectaries in the field. The sugar concentration of the nectar can be determined in the field with hand-held refractometers (e.g. (Corbet, 2003) or, when the nectar is secreted in small quantities, alternative methods may be required (e.g. (Kearns and Inouye, 1993); (Dafni et al., 2005)). The sample size should not be fewer than 20 randomly selected individual plants at each field. Ideally,

the time of sampling should be fitted to the intensive foraging period of the bees (i.e. ideal meteorological conditions and when considerable foraging activity is observed). The samples obtained from the same field can be averaged. This sampling regime will result in at least 25 average data, which than can directly be used for refined (Tier 2) shortcut value calculations. It is, however noted, that the data available in the data base for Tier 1 (Annex C of the Supplementary Document) can be reused, that might reduce the number of varieties/locations to be involved in the experimental data collection for this refinement step.

As regards the adjacent crop scenario and the succeeding crop scenario, the same principle applies (i.e. the SN parameter could be refined). However, any refinement for those scenarios become practical only if the crops of those scenarios are well defined (for which, no detailed guidance can be given). Otherwise, as in Tier 1, crops belonging to sugar category 1 will be considered as default. No guidance could be given for the definition of the crops that might belong to the adjacent crop scenario and the succeeding crop scenario.

The composition of the plants of the non-crop scenarios (weeds in the field and plants at the field margin) vary in space and time. Therefore, any experimentally established value (obtained from a certain landscape at a certain time) might not be considered as sufficiently representative for a risk assessment.

Further details are available in Appendix S of EFSA (2013) and for general guidance for exposure field studies, please consult Section 5.3.8.

5.5.7 Landscape factor (LF_{po}, LF_{ne})

The Landscape Factor (LF) considers that bees will not forage solely on a treated crop, but will also visit other flower sources in the landscape. The resulting exposure for the PPP under evaluation depends on the foraging behaviour of the bee species, the crop and the landscape characteristics. A LF of 1 for the treated crop scenario means that 100% of the food entering the hive/nest originates from treated fields of that crop.

In Tier 1, the residue levels are based on residues collected from flowers only, and a LF <1 is recommended for the chronic dietary exposure estimation in pollen for honey bee adults and larvae. This LF in Tier 1 is not a single value, but a range of values (see in 5.3.7, above) that feeds into the Monte Carlo exposure estimation (i.e. into the shortcut values).

A refined LF can only be used when refined residue values are not determined from field studies with bees, because in that case, it is already implicitly included. Therefore, in case refined residue values from field studies with bees are used, Tier 2 LF should be set to 1. In case when Tier 1 residue values, refined residue values from semi-field studies or from field studies without bees are used (i.e. samples taken directly from the flower), refinement of the LF is possible under certain conditions. This allows a crop-specific refinements.

Refinement of the LF is only advisable for honey bees. Other bee species can theoretically be used, but extrapolation to their respective groups will be very challenging. It would have to be justified why the foraging behaviour of the studied bumble bees or solitary bee covers the foraging behaviour in that specific landscape and crop covers all other species in those groups.

Also, no protocol for nectar LF could currently be recommended by the WG. This is because there is very limited experience with flower source determination from nectar and no data were available to the WG (see Annex A of the Supplementary Document).

General recommendations for LF refinement studies are given below. These align, where relevant, with the requirements for residue studies with bees from open field conditions as described under 5.5.8 and in Annex B of this document and take into account the findings in the dataset considered for the Tier 1 LF (see also Section 5.3.7 of the Supplementary Document).

The representative crop should be used, or a justification should be provided for extrapolation between crops (see 5.5.8, below).

Depending on whether the landscape can be characterized as 'minimal alternative forage' or ' randomly selected', hives at 5 or 15 locations needed to be used, in line with Section 5.5.8. One hive per location is sufficient, but it may be advisable to use more for backup.

LF refinement studies can be performed in untreated fields and can thus be used for any PPP provided that the other study conditions such as the landscape range are representative for the proposed use. It is noted that in case a PPP would have a repellent effect, using generic studies in untreated fields represents a worst case. However, if a PPP has an attractant effect (not excluded phenomena based on the dataset collected for Tier 1, (see Section 5.3.7 of the Supplementary Document) the exposure could be underestimated. If there are indications from other research that a PPP has an attractant effect, generic studies should not be accepted.

The analysis in Section 5.3.7 of the Supplementary Document shows that the percentages crop pollen collected differs between different sampling dates within a field. Sampling should be performed at least three sampling times during the flowering period to ascertain that the most attractive crop phase is caught, in line with in line with Section 5.5.8.

Pollen can be collected with pollen traps, placed in front of the hive for ca. 4 hours during the time of day that the bees are most active. See in Section 5.5.8 for more guidance.

The plant species origin of the pollen can be determined with palynology (microscopic pollen analysis) or molecular methods (DNA metabarcoding). A description of these methods is given in 3.1.1 and 3.1.2 in Delaplane et al. (2013). Determination by colour is not accepted as this is not an accurate method. The methods used must be described in detail in the study report.

Data analysis will lead to a percentage of crop pollen collected at three timepoints in five or fifteen locations, i.e. 15 or 45 data points. Per location, the highest percentage is taken for the assessments for chronic adult and for the larvae, which results in a dataset of 5 or 15 values (which may have been collected at different sampling times).

If the studied crop was oilseed rape or *Phacelia*, these data should be added to the dataset that is already available (see in Section 5.3.7 of the Supplementary Document) and used for the revised shortcut value calculation.

For any other crop, the 5 or 15 datapoints replace the Tier 1 LF dataset to calculate revised shortcut values (chronic adult and larvae assessment).

The dataset underlying the Tier 1 LF contained some cases of 100% crop pollen collection on single sampling dates (see 5.3.7). Therefore, the data set of landscape factors for the Tier 1 exposure assessments was only applied for chronic assessments, but not for the acute assessment (i.e. Tier 1 LF for acute assessment is 1). However, if crop-specific data are available, LF refinement is possible even for the acute assessment. In this case, the highest of the 5 or 15 datapoints has to be used.

Modelling

An alternative to measure the concentrations in nectar and pollen entering the hive in field studies is to model the foraging behaviour on all attractive plants, including the treated crop, in the foraging area at the landscape level. However, at the time of the writing this document, the WG cannot recommend a specific model. Nevertheless, modelling such parameters might become a reasonable tool in the near future.

5.5.8 Residue Unit Dose (RUD)

The generic RUD values used for the Tier 1 exposure estimations were established in a conservative way. PPP and crop specific or even GAP specific RUDs (i.e. concentration levels in pollen and nectar) can be significantly different from the generic RUDs. Therefore, using more specific RUDs in the exposure assessment could lead to more realistic risk estimation. PPP/crop/GAP specific RUDs can be derived from field measurements. Appendix G of EFSA (2013) includes some guidance on how such field measurements might be conducted. The WG has reviewed the recommendations and the requirements for those field measurements, which are also called as exposure field studies. The detailed guidance for performing such exposure field studies are included in Annex B of this Guidance Document.

For a Tier 2 exposure estimation, the concentrations measured in the exposure studies have to be expressed as RUDs considering 1 kg/ha.

A minimum of five reliable RUDs from five different locations in the area of use of the substance are needed if the RUDs are derived from residue levels measured in pollen and nectar collected directly from flowers and/or collected from foraging bees confined to tunnels/cages and/or collected from foraging bees in open field where minimal alternative food sources are available in the landscape. In case the locations are randomly selected ("randomly selected landscape" field studies), the minimum number of reliable RUDs necessary to refine the exposure for a specific GAP is fifteen (see Annex B).

The RUDs should be generated in a way that they reflect exactly the representative use (specific GAP) of the PPP under assessment, i.e. the directly treated crop needed to be the crop in the GAP. However, the RUDs from a particular exposure field study that follows a certain GAP might be used for a risk assessment for another GAP if the circumstances in which those RUDs were generated are likely equally severe or represent a worst-case situation of the other GAP (exposure envelope approach). In general, the following aspects might be considered for this assessment:

- RUDs derived from field residue studies with application(s) during the flowering of the focal crop (BBCH 60 – 69) are considered to cover also GAPs with application(s) before flowering of the focal crop
- RUDs derived from field residue studies with downward spray application to highly attractive model crops such as *Brassica napus* (oilseed rape) and/or *Phacelia tanacetifolia* and residues are bee-collected, are consider covering other GAPs with different crops and the same application methodology
- RUDs derived from residues collected directly from flower of the focal crop are comparable to RUDs derived from residues collected from bees in semi-field studies, and both of these types of RUDs are considered to cover residues derived from "randomly selected landscape" field studies (Section 1.1 of Annex C –)
- If a risk envelope approach is used, RUDs derived from field studies where the intended crop for the proposed item is not used, could be considered acceptable providing that

background information as well as the rationale developed in proposing the worst-case GAP are clearly explained and reported

- If RUD values are available on both bee pollen/nectar and plant pollen/nectar from the same semi-field study, the values from bees only should be considered

The maximum observed in the available samples should be retained as representative of the exposure in each particular field experiment. This does not imply that the overall risk assessment has to be regarded as overly conservative, since the sampling frequency pattern in the studies does not guarantee that the actual maximum occurrence was picked up by the maximum measured in the samples taken. Nevertheless, it is expected that the assessment based on these principles may still be considered to represent a realistic worst-case exposure for the different substances and uses assessed.

For the refinement of the treated crop scenario, extrapolation between RUDs from different crops is inappropriate when residual pesticides in soil are known to be taken up by crop roots and subsequently translocated to aboveground tissues. This is because the different physiology of different crops, including the time from emergence to flowering leads to different translocation and levels of residues in different crops.

5.5.9 Half-life in pollen and nectar ($DT50_{po}$, $DT50_{ne}$)

The generic DT_{50} values used for the Tier 1 exposure estimations were established in a conservative way. Active substance and crop specific decline data can be different from the generic Tier 1 estimations. Therefore, using more specific residue decline information in the exposure estimation could lead to more realistic risk assessment. The required specific DT_{50} values can be derived from field measurements. Pertinent requirements for those field measurements have been set by the WG. The detailed guidance for performing such field measurements are included in Annex A.

It should be noted that this parameter has an effect only on certain chronic cases (i.e. for adult chronic assessments and the larval assessment for honey bees).

5.5.10 Half-life in plant matrixes ($DT50_{pnt}$)

The generic DT_{50} value used for the Tier 1 exposure estimations were established in a conservative way. Active substance and crop specific decline data can be different from the generic Tier 1 estimations. Therefore, using more specific residue decline information in the exposure estimation could lead to more realistic risk assessment. The required specific DT_{50} values can be derived from field measurements. Pertinent requirements for those field measurements have been set by the WG. The detailed guidance for performing such field measurements are included in Annex A.

It should be noted that this parameter has an effect only on pre-flowering spray and granular applications. The refined DT_{50} value could be used in place of the default $DT50_{pnt}$ parameter value, which is relevant only for multiple applications made before the flowering. In addition, – only for pre-flowering spray applications- the $DT50_{pnt}$ plays a role in the PFF parameter, which could then be also refined (see in Section 5.5.3).

5.5.11 Number of applications (n_{de}, n_{du})

The number of applications is defined by the GAP of the PPP for which the risk assessment is performed. Therefore, no refinement option is possible.

5.5.12 Interval between multiple applications ($i_{be,}$ i_{du})

The interval between applications is defined by the GAP of the PPP for which the risk assessment is performed. Therefore, no refinement option is possible.

5.5.13 Time window (w)

In a number of risk cases the residue decline during the exposure phase (i.e. flowering time) does not play a role (Tier 1 parameter value of w is set to 1). For example, all acute risk assessments, by default, consider a single peak exposure. This does not change between the Tiers, therefore no refinement is possible. In theory, the residue decline over time can be considered for all chronic assessments; some of the Tier 1 values are set for a certain window (5 or 10 days), while for other chronic cases the conservative value of 1 was set. In any case, refinement of the Tier 1 parameter value w (tough theoretically possible) is not recommended for different reasons, which are explained below.

In case of all pre-flowering assessments no consideration for residue decline (therefore no w > 1) is proposed for Tier 1 because an even residue translocation rate from the plant tissues towards pollen and nectar is assumed. This assumption is based on the limited data that is available for those situations. The WG proposes further research on this area in order to better understand the residue behaviour. Once more data are available, the parameter value might be reviewed. Similarly, no consideration for residue decline is proposed for Tier 1 for solid formulations (seed treatment, granules). This is again due to lack of understanding the residue behaviour from these formulations. Once more data are available, the parameter value might be reviewed for those situations.

No consideration for residue decline is proposed for Tier 1 for bumble bee and solitary bee larva. This is because the risk assessment for these larvae is meant to cover all bumble bee and solitary bee larva. In order to set a time window over which considering the residue decline would result in a sensible risk assessment, further information would be needed (see in 5.3.13). Some information is available, but only for some species of the bee groups (including the test species). Using a refined w parameter in a Tier 2 risk assessment would result in a specific risk assessment relevant only for those species for which the refined w parameter is relevant. However, further extrapolation to other species was not reasonable. Therefore, until better information is available (e.g. more knowledge on the length of developmental period of the relevant species), refinement of the parameter value w is not recommended.

For honey bee larvae, the Tier 1 parameter value is realistic, as it was set based on relevant biological information (e.g. brood development). Therefore, no further refinement is possible.

For adult bees (all bee groups), the Tier 1 parameter value was set in way that it represents a significant period from the active lifespan. In addition, it fits with the available testing methods. Overall, it is considered to be realistic for a chronic risk assessment for most of the PPPs. For PPPs prone to time reinforced toxicity, a specific risk assessment scheme is proposed that considers different lifespan (see Chapter 6).

5.5.14 Shortcut values ($SV_{po,soil}$, $SV_{ne,soil}$)

The following parameters contributing to the shortcut values might be refined: LF, SN, PEC_{pw}. In case of appropriate Tier 2 data are available for one or more of those parameters (see corresponding sections of this Chapter), the SV should be recalculated by considering the Tier 2 data. The other, not refined parameters, must take the Tier 1 parameter value. The resulting values are considered to be Tier 2 shortcut values. The method of calculation is the same as described in 5.3.14.

5.5.15 Predicted concentrations in soil pore water (PEC_{pw})

In the current EU regulatory framework predicted concentrations of active substances and their degradation products in soil (PECsoil) are calculated based on simple models with the realistic worst case DT_{50} (in many cases this will be longest field dissipation DT_{50}) and a fixed "soil scenario" (soil with a bulk density of 1.5 g/cm^3 , and a mixing depth of 5 cm for applications to the soil surface or 20cm where incorporation is involved). The PECsoil are expressed as concentration on a dry weight basis (mg/kg), while pore water concentrations (mg/L) are not used in standard risk assessments for soil organisms. However, the pore water concentration is regarded as the bioavailable fraction for plant root uptake as soil water serves as the carrier to move the chemicals into plants (see Chapter 2 of Annex I of the Supplementary Document).

At the time of writing this guidance, the new EFSA Guidance on soil exposure (EFSA, 2017) and the supporting software tools (PERSAM, PEARL and PELMO) are under the scrutiny of the SCOPAFF (Standing Committee on Plants, Animals, Food and Feed) for noting. In line with the exposure assessment procedures for the other environmental compartments ((FOCUS, 2001, European Commission, 2014)) a tiered approach is recommended. The first tier is based on simulations for one predefined scenario per regulatory zone carried out with the simple analytical model PERSAM and the scenarios apply to both annual and permanent crops. In case the exposure assessment is focussed on a specific crop, a spatially distributed version of PERSAM (at Tier-2) is used and the 95th percentile concentration considering all agricultural fields within a regulatory zone will be assessed. The model is applied to all 1 km by 1 km grid cells (combination of soil moisture content, soil bulk density, soil organic matter, temperature) where the target crop is present. Should the assessment at Tier-2 still indicate an unacceptable risk, there is the option to move to Tier-3A, which is a more realistic approach based on the area of a given crop but uses numerical models (PEARL and PELMO) which consider also the dissipation at the crop canopy and the wash-off processes. In all tiers, the concentration in total soil and the concentrations in soil pore water averaged over various depth and time windows are calculated.

As Tier-3A calculations requires more substance-specific input values and slightly more efforts than the use of the PERSAM tool, The WG suggests starting from Tier 2 calculations with PERSAM to derive the predicted soil pore water concentrations (PEC_{pw}) to refine the dietary exposure assessment for through soil contamination (see paragraph 5.1.2.2). PEC_{pw} should be derived over 20 cm soil depth and at the time of flowering of the target crop using median or average substance properties from the dossiers. For the succeeding crop scenario, as conservative assumptions for the time of the beginning of the flowering succeeding crop, PEC_{pw} should be calculated after 75 days from the last application on annual double crops, after 120 days after the last application on permanent crops, and 150 days after the last application on annual crops (Annex I of the Supplementary Document). This is applicable to spray applications as well as solid applications to the treated crop. A detailed description of input screens and the Tier-1 and Tier-2 calculations of the PERSAM v3.0.5 can be found in the user manual (NV, 2021).

5.6 Exposure assessment for the dietary screening step

As described above in this Chapter, the exposure assessment for the dietary route of exposure is a complex process, involving several steps and numerous parameters. Therefore, the WG has formed a simplified method for the deviation of PEQdi values. This screening PEQdi can be used in the combined risk assessment (see Chapter 7). Applying this screening is an option, but not mandatory. The method for PEQdi calculations for the screening step is a simplified version of the Tier 1 method as described in Section 5.1.2, resulting in conservative exposure estimations compared to Tier 1. In the simplified model the application rate (AR; g/ha) and the number of applications (n) - as described in Section 5.1.2 - has to be combined with a factor in the following way:

$$PEQdi = AR\ n\ B$$

where B is a factor that depends on the category and the application method as presented in Table 20, below.

Table 20: The values of factor B for each category by application methods

Category for the risk assessment	Risk cases	Constant B to be used for downward spray and granular application (µg/bee or µg/bee/day or µg/larva/developmental period)	Constant B to be used for sideward/upward spray application (µg/bee or µg/bee/day or µg/larva/developmental period)	Constant B to be used for seed treatment (µg/bee or µg/bee/day or µg/larva/developmental period)
Honey bee	*acute adult*	6.4	9.0	1.08
	chronic adult	6.2	9.0	1.04
	larva	7.2	9.2	1.21
Bumble bee	acute adult	10	13.7	1.67
	chronic adult	9.6	13.3	1.62
	Bombus terrestris larva	33.7	48.5	5.66
Solitary bee	acute adult	0.70	0.94	0.12
	chronic adult	0.67	0.90	0.11
	Osmia bicornis larva	38.2	57.5	6.41
	Osmia cornuta larva	47.2	68.8	7.93

Due to the simplifications applied, this method results in more conservative PEQdi values than any of the Tier 1 scenarios. Therefore, a single calculation will cover all the five Tier 1 scenarios. Further details on the considered simplifications, the derivation of the exposure model for the screening step and the derivation of the constant values are reported in Section 5.6 of the Supplementary Document. No screening step is proposed for the contact exposure.

6 Effect assessment in lower tiers

As explained in Section 3, the aim of the effect assessment is to identify the relevant hazard parameters or 'effect endpoints' to be used together with the PEQ_j for estimating the levels of risk.

At lower tier level, the hazard is defined by the dose-response curve described by the median lethal response dose (hereafter $LD_{50,j}$) and the slope (hereafter $slope_j$, i.e. the parameter describing the steepness of the dose-response curve) investigated in standard laboratory tests. Once again, the suffix *j* indicates one of the four risk cases (i.e. acute oral, acute contact, chronic, larvae)

The WG considered that the use of the full dose-response relationship would improve upon the previous methodology, which relied on single effect point estimates (e.g. NOED, LD_{50}) without an explicit consideration of the predicted level of effect triggered by a different exposure level. Therefore, the WG propose defining the effect endpoints as function of both $LD_{50,j}$ and $slope_j$, determined by a log-logistic dose-response function.

The *a-priori* selection of a specific dose-response model represents a deviation from the recommendation of the BMD Guidance Document (EFSA, 2017) to perform model averaging. Nevertheless: 1) the log-logistic models are by far the most commonly used models for describing dose–response data ((Ritz, 2010); 2) a comprehensive analysis of many bee-specific datasets (see Section 5.2 in the Supplementary document) confirmed that this model fit the data well in most cases; 3) selecting a specific model whose parameters have a straightforward interpretation presents clear advantages in terms of communication. Therefore, the WG considered this deviation acceptable.

In order to use appropriate hazard parameters (i.e. $LD_{50,j}$ and $slope_j$) in risk assessment, risk assessors should carefully consider various aspects, which are presented in the following sections. These are:

- Definition of hazard parameters in experimental studies indicated by the legal requirements (Section 6.1);
- Combination of hazard parameters from equivalent studies performed with the same substance and the same species (Section 6.2);
- Derivation of a surrogate dose-response beyond the tested range (Section 6.3);
- Consideration of time-reinforced toxicity (Section 6.4);
- Extrapolation of the hazard parameters between species (Section 6.5).

6.1 Definition of hazard parameters in experimental studies indicated by the legal requirements

6.1.1 Legal requirements

Under Regulation (EC) No. 1107/2009, the data requirements are set in Regulation (EC) No. 283/2013 (for active substances) and Regulation (EC) No. 284/2013 (for PPPs), for bees they are described in point 8.3 and 10.3, respectively.

Any dossier should contain the tests that are required for that particular use (see Chapter 4), and these tests should be performed according to the standard guidelines (OECD test guidelines, or existing protocols (pending validation and adoption of new test guidelines) see overview in Table 21. In addition, information that is relevant for endpoint setting may come from public literature and non-guideline studies. The relevance and reliability of all available studies should be considered for the overall selection of endpoints.

More specifically, Regulation (EC) No. 283/2013 requires that:

- Acute oral and contact toxicity to bees is tested where bees are likely to be exposed (8.3.1.1);
- Chronic toxicity to bees is tested where bees are likely to be exposed (8.3.1.2)
- Honey bee brood study shall be conducted to determine effects on honey bee development and brood activity, unless exposure of brood is not possible. The bee brood test shall provide sufficient information to evaluate possible risks from the active substance and the plant protection product on honey bee larvae (8.3.1.3)

6.1.2 Toxicity studies

Many OECD guidelines and guidance documents for bee testing have become available in recent years or are under development (see Table 21).

Table 21: Overview of the current availability of standard test guidelines

Test type	Honey bee	Bumble bee	Solitary bee
Acute oral test	OECD Test Guideline No. 213 (OECD, 1998a)	OECD Test Guideline No. 247(OECD, 2017a)	Standard test methods not yet available Draft version available for *Osmia* species (Roessink I, 2019)
Acute contact test	OECD Test Guideline No. 214(OECD, 1998b)	OECD Test Guideline No. 246 (OECD, 2017b)	Standard test methods not yet available Draft version available for *Osmia* species (Roessink I, 2017)
Chronic test	OECD Test Guideline No. 245 (OECD, 2017c)	Standard test methods not yet available	Standard test methods not yet available
Test on larvae	OECD Guidance Document No 239 (Repeated Exposure)	Standard test methods not yet available	Standard test methods not yet available

For all standard tests on adult honey bees and bumble bees (ring tested on *Bombus terrestris* and *Bombus impatiens*), the main endpoint is mortality. However, sublethal effects are also recorded (see Chapter 9).

Though not required in all test guidelines, the WG highly recommends to always perform analytical verification of the exposure during the test.

The declared purpose of all these standard test guidelines is the derivation of a reliable LD_{50} expressed as mass of a.s.·bee^{-1} (a LDD_{50} expressed as mass of a.s.·bee^{-1}·day^{-1} for chronic tests). Generally, the range of concentrations used in the tests should be wide enough to yield mortality effects that spans from very low to high (i.e. >50%), in order to derive a reliable LD_{50}. This is often enough to derive a reliable slope of the dose-response as well.

The OECD 'unclassified' Guidance Document No. 239 (OECD, 2016) is available to address the data requirement for honey bee brood. The test exposes young honey bee larvae in artificial cells for four days via treated food in/on which they lie and follows them up to and including the moment of adult emergence from the pupae. The main recorded endpoints are mortality at larval and pupal stage and emergence, which can also be considered a proxy for lethal effects.

Guidance document 239 should be used instead of an earlier publication, OECD guideline 237 (OECD, 2013), which only exposes the larvae for a single day and ends at pupation.

Test recommendations for larvae other than the honey bee are not available and the current data requirements specifically address only honey bees. In the relevant OECD Guidance Document, it is specified that "*The method aims at the determination of a No Observed Effect Concentration/Dose (NOEC/NOED) and, if data allows, EC50/ED50*". Indeed, even in the EFSA (2013), the focus was on deriving a NOED. To fulfil the requirement of the updated risk assessment it is recommended that any larval test should be conducted with doses able to cover effects at least greater than 50%, so that a reliable LD_{50} (expressed as mass of a.s.·larva^{-1}·developmental period^{-1}) and a reliable slope are derived. Note that this does not entail any modification of the existing OECD Guidance Document, which already foresees this possibility.

Irrespective of the type of test, in some circumstances, the experimental data may show no or very low mortality at one dose and (almost) full mortality at the next one, especially when the dose-response is very steep. This often leads to a high uncertainty[8] around the mean value of $slope_j$ parameter. In such cases, it is important to assess whether the predicted exposure PEQ_j happens to be in the steep part of the dose-response: if this is the case, it is appropriate to use the 95% lower limit of the uncertainty range as representative of the $slope_j$ parameter. This ensures that a worst-case effect is predicted for doses below the $LD_{50,j}$. Otherwise, this additional step is not necessary.

For experiments carried out as limit tests, a study-specific slope cannot be obtained, and a default value should be used (see Section 6.3).

It is acknowledged that tests submitted in dossiers may have been performed using a draft version of a guideline that has since been superseded by a final version. It is up to the risk assessor to determine whether such test is still acceptable for use in the risk assessment or if a new test should be performed. This will be determined on a case-by-case basis and may depend among other on:

- the timing of the evaluation compared to the guideline update, acknowledging that studies are generally performed before submission of a dossier;
- the changes made to the specific guideline: are the deviations from the final guideline expected to have severely impacted the outcome of the test?; and
- the importance of the endpoint in the overall risk assessment, e.g. are higher tier tests available that cover the first tier endpoint?

Similarly, the reliability of endpoints from draft test guidelines which have not been finalised by the OECD should be carefully considered.

6.1.3 Active substances and Plant Protection Products

According to Regulation (EU) 284/2013, when the toxicity of the PPP cannot be reliably predicted from the active substance, studies (acute oral and contact, chronic, honey bee brood) performed with the PPP are required (from 10.3.1.1 to 10.3.1.3). In the case of mixture i.e. PPP where the active substance is always intended to be used together with a safener and/or synergist and/or in conjunction with other active substances, toxicity of the mixture cannot be predicted based on the data of the active substance, therefore data on PPP-mixture are always required (see Table 22). For the risk assessment of PPP containing more than one active substance see Chapter 12.

As further interpretation of the legal requirements, for PPP containing only one active substance, at least acute (contact and oral) studies are required for both the active substance and the PPP, as this is the basis for a toxicity comparison between the active substance and the PPP. A chronic toxicity study and a honey bee brood study with the PPP can be waived, if based on the comparison between acute toxicity studies, the PPP results in a comparable or lower toxicity than

[8] As a tentative rule of thumb, the slope is considered highly uncertain if the ratio between the 95% upper limit and the 95% lower limit of the uncertainty range is greater than 10. This may need to be updated when more experience is gained.

the active substance. A ratio of 3 should be used for the investigation of potential higher toxicity of the PPP based on the acute toxicity endpoints. Therefore, if:

- $LD_{50,acute}$ (a.s.) / $LD_{50,acute}$ (PPP^{9}) > 3 → all the data requirements must be fulfilled by providing, in addition to the acute, also chronic and brood data for both the active substance and the PPP.
- $LD_{50,acute}$ (a.s.) / $LD_{50,acute}$ (PPP^{9}) ≤3→ no further data on PPP are needed.

For considerations concerning (1) availability of several equivalent tests carried out with the same species and the same test item and (2) uncensored LD_{50} values, please see Sections 6.2 and 6.3, respectively. Further considerations about the comparison between active substance and representative PPP are reported in Section 6.6.1.

Table 22: summary of the data requirements for the active substance and the formulated products

Tier 1 study type	**Study with active substance (where bees are likely to be exposed) required?**	**Study with formulation required?**	
		Formulation with one active substance	**Formulation with more than one active substance**
Acute oral	Yes	Yes[a]	Yes
Acute contact	Yes	Yes[a]	Yes
Chronic oral toxicity to adults	Yes[c]	Pending on the comparison between acute studies[b]	Yes
Toxicity to larvae	Yes[c]	Pending on the comparison between acute studies[b]	Yes

[a] Acute studies with the formulation can be waived when the toxicity can be predicted on the basis of the active substance (e.g. when the formulation consists of the active substance only, or of the active substance in water).

[b] Generally, a study with the active substance will be sufficient; however, if there is an indication from the acute oral study that the formulation is more toxic than the active substance, then the formulation should be tested. In determining whether there is a difference then the endpoints should be expressed in terms of active substance. If the acute formulation endpoint expressed as active substance **is more toxic by at least a factor of 3** than the acute endpoint for the active substance, then it can be assumed that the formulation is of greater toxicity and hence chronic and larval testing should also be carried out using the formulation. If the difference is less than a factor of 3, then testing adult chronic and larval toxicity with the active substance is sufficient.

[c] In case of poorly soluble substance, a single study on the formulated product might also be appropriate as surrogate if higher solubility levels are expected with the formulated product under the test conditions.

6.2 Combining equivalent studies performed with the same test item and the same species

Often, multiple equivalent tests are available e.g. more acute contact tests with honey bees and a certain test item. Differences between the results of these tests are likely driven by experimental variability, thus averaging is generally proposed.

[9] Expressed in terms of a.s.

It must be noted that this averaging procedure is not generally recommended between tests carried out with different test items (e.g. one active substance and a formulation). On the contrary, if multiple equivalent tests are available with the active substance and/or with a formulation, averaging among tests performed with the same test item is recommended before checking whether the formulation shows higher toxicity compared to the active substance (see Sections 6.1 and 6.6)

Whenever multiple equivalent studies are available, it is recommended that the representative dose-response makes use of the geometric mean of the $LD_{50,j}$ derived from the individual studies. This recommendation is in line with previous Guidance Documents (e.g. (EFSA, 2009).

However, in presence of right-censored values, more considerations are needed:

- If one test only allows the derivation of a defined (uncensored) LD_{50}, such value should be used in the representative dose-response for the risk assessment. In this case, right-censored values are simply disregarded.
- If more tests allow the derivation of a defined LD_{50}, the geometric mean of those should be calculated and used in the representative dose-response for the risk assessment. In this case, right-censored values are disregarded.
- If all available tests resulted in right-censored LD_{50}s, the highest should be considered for extrapolation according to the procedure detailed under Section 6.3.

When drafting this Guidance Document, no information was available concerning the distribution of slopes from fully equivalent studies performed on the same test item. As a pragmatic solution, it is proposed that the $slope_j$ parameter of the dose-response to be used in the risk assessment should make use of the geometric mean of all reliable slopes derived from the individual studies. This is a conservative approach, since (1) the geometric mean will never be higher than the arithmetic mean and (2) because shallower slopes represent a worst-case (higher predicted effects) for doses below the LD_{50}.

In line with the recommendations given for the $LD_{50,j}$, whenever the slope from a certain test is undefined (e.g. limit test), this should not be considered in deriving the representative $slope_j$ parameter for the dose-response used in the risk assessment.

6.3 Derivation of a surrogate dose-response beyond the tested range

For substances with low toxicity and 'difficult' substances (e.g. with low solubility), it is often the case that the highest tested dose or the 'limit dose' is expressing an effect <50%. In this event, the LD_{50} is often referred as right-censored value (e.g. LD_{50} > 100 µg a.s./bee). In the past, the lower bound of the endpoint was used in the risk assessment. This approach overestimates the actual toxicity and might bias the risk assessment of substances with low toxicity.

In these cases, the experimental data do not allow describing a full dose-response. In fact, often neither an $LD_{50,j}$ nor a $slope_j$ is reliably estimated. However, a surrogate dose-response can still be derived by making some conservative assumptions.

For any specific dose d, assuming a shallower slope will determine a prediction of higher effects for any dose < d. In the risk assessment scheme proposed in this Guidance Document, any predicted exposure PEQ_j equivalent to a dose causing effects >10% at the individual level will immediately trigger a high risk at the tier-1 (see Chapter 7). This means that underpredicting effects above 10% has no consequences on the outcome of the risk assessment. On the contrary, it is of utmost importance to ensure that a conservative approach is taken when predicting effects

in the range 0-10%, as their combination will determine the outcome of the risk assessment (see Chapter 7).

As a conservative approach, a default $slope_j$ of 1.43 can be used whenever a specific value cannot be reliably determined from the experimental data. This corresponds to the 10th percentile of the slope distribution based on an analysis of dose-response curves obtained from a large number of substances (see Section 6.3 in the Supplementary document).

Once $slope_j$ is fixed, a surrogate $LD_{50,j}$ can be derived by multiplying the highest (or single) tested dose to an appropriate extrapolation factor, depending on the observed effect. For sake of simplicity and conservativeness, the observed effects are discretised in four effect intervals, with the higher end being used for deriving the extrapolation factors.

As significant differences among slopes were not recorded between groups of substances and test types, a unique set of extrapolation factors (reported in Table 23) can be applied to all kind of tests.

Table 23: Extrapolation factors for different intervals of effect.

Effect observed at the highest tested dose	**<10% effect**	**10-20% effect**	**20-30% effect**	**30-40% effect**
Extrapolation factor to be applied to the highest tested dose	4.6	2.6	1.8	1.3

To summarise, whenever a full dose-response cannot be derived from the experimental data, a surrogate dose-response can be derived by:

- Using a worst-case default $slope_j = 1.43$
- Using the appropriate extrapolation factor to derive a surrogate $LD_{50,j}$.

6.4 Time-reinforced toxicity (TRT)

A substance shows time-reinforced toxicity when its toxic effects from exposure to low doses for a long period of time are higher compared to effects from exposure to higher doses for a short period of time (i.e. its toxic effects are reinforced by exposure time). Note that this phenomenon was called 'accumulative toxicity' in the EFSA (2013).

The studies required by Regulation (EC) No. 283/2013, 8.3.1 and Regulation (EC) No. 284/2013, 10.3.1 are limited to studies assessing the effects from fixed exposure durations (acute and chronic). This approach is valid when toxicity is mainly dependent on the dose. However, in case of substances which show time-reinforced toxicity, the impact of low doses may be underestimated if the exposure period tested in the laboratory is shorter than the environmentally relevant length of exposure. Therefore, the WG considered it necessary to always assess whether a substance shows time-reinforced toxicity.

The assessment strategy for TRT included in this guidance document (see Chapter 8) is based on the results of the honey bee chronic toxicity study according to OECD guideline 245 (OECD, 2017c). Therefore, in principle, no additional studies need to be performed for the TRT assessment.

The ten-day duration of a chronic study according to OECD 245 has been criticized as being too short to address longer exposure to low doses. However, an analysis performed by the WG has shown that data from a ten-day study can be used to reliably predict the toxicity for a longer

exposure period (please refer to Section 5 of Annex G to the Supplementary Document for details). It should however be noted that the observed mortality in the study should be sufficiently high for it to be useful for the TRT assessment (i.e. it should be possible to calculate an LDD_x value for at least 4 days; refer to Chapter 8 for further details). Therefore, the doses tested should be carefully considered.

In case a conclusion on the occurrence of TRT cannot be drawn based on the available chronic toxicity study, a study specifically designed to investigate TRT can be performed. For the design of such a study, there are different options (which can also be combined), as detailed in Chapter 8.

Whenever a substance shows TRT behaviour, the lifespan dose-response obtained from the TRT assessment substitutes the 10-days dose-response obtained directly from the chronic testing with honey bees.

6.5 Extrapolation between species

The lack of toxicity data for bumble bees and solitary bees makes it difficult to assess the risk of pesticides for these bee groups. The issue of how LD_{50}s differ among bee species has been investigated from different perspectives, in order to get suitable extrapolation factors. Details about the analysis underlying this section are available under Section 6.5 of the Supplementary document.

The available data allowed a meaningful comparison of acute, mostly contact, tests, nevertheless this was not the case for other type of studies and duration due to the lack of standardised chronic and larvae tests for species other than *Apis mellifera*. Therefore, in absence of suitable alternatives, and until better data become available, the WG recommends that the derived factors are applied to all types of tests with adult bees, despite being uniquely derived from the acute (mainly contact) tests.

The availability of bee weight measured in some ecotoxicity experiments allowed establishing a generic (substance-independent) relationship between LD_{50} and bee weights. Thus, collection of adult bee weight data for a representative number of European bee species (~10%), allowed calculating extrapolation factors from standard species to smaller bumble bees and solitary bees, in order to protect at least 95% of European species with 95% confidence. The WG has considered this to be a very conservative approach, that can be revised if more information will become available These factors are referred to as 'Toxicity extrapolation factors' (*Tef*) and are reported in Table 24 for standard species.

The *Tef* values only refer to the hazard definition, so that smaller bees would be characterised by smaller LD_{50} (higher hazard). However, as explained in Chapter 5, it is noted that exposure estimates are also dependent on body weights (see Section 5.3.5) and body surface (see Section 5.2.3), so that smaller bees are characterised by lower exposure.

For larvae, no suitable information is available to relate neither the LD_{50} nor the predicted exposure levels to the bee size, so an extrapolation equivalent to the ones performed for adult bees is not possible. As explained in Chapter 5, exposure estimates are based on *Bombus terrestris* for bumble bees and *Osmia* species (*O. rufa* and *O. cornuta*) for solitary bees. None of these species are significantly smaller than honey bees. As such, a *Tef* = 1 is proposed from honey bee larvae to bumble bees and solitary bee species (see Table 3).

Table 24: Toxicity extrapolation factors (*Tef*). Standard $LD_{50,j}$ should be divided by these factors to obtain an estimate of an $LD_{50,j}$ protective of 95% of the species in the group.

Category - extrapolation from -	***Tef* for extrapolation to**	
	5th percentile BB weight	5th percentile SB weight
Standard honey bee adult (*A. mellifera* worker)	2.8	228
Standard bumble bee adult (*B. terrestris*[a] worker)	7.8	-
Standard solitary bee adult (*O. rufa* ♀)	-	191
(O. cornuta ♀)	-	408
Standard honey bee larva (*A. mellifera* worker)	1.0 [b]	1.0 [b]

[a] OECD test guidelines No. 246 and 247 were also ring tested with *B. impatiens*. If data are available with this species, both *Tef* and food consumption values should be recalculated based on the appropriate body weight. For *Tef*, the formula is available in the Supplementary document under Section 6.5.4. Camp et al. (2020)reported an average weight of 178 mg for *B. impatiens*.

[b] *Tef* not meant to address the 5th percentile species in terms of weight, but rather *Bombus terrestris* for bumble bees and *Osmia* species for solitary bees, i.e. species used to estimate the exposure levels to bumble bees and solitary bees.

Hence, the appropriate extrapolated $LD_{50,j}$ for bumble bees and solitary bees will be derived from the following equation:

$$Extrapolated\ LD_{50,j} = \frac{(surrogate)\ Standard\ LD_{50,j}}{Tef}$$

In most of the cases, the standard $LD_{50,j}$ (which can, in some cases, be a surrogate $LD_{50,j}$) is derived for honey bees. Nevertheless, if data are available on other standard species, those should be used in the derivation of the extrapolated $LD_{50,j}$ for their specific bee group.

The extrapolation factors presented in this section are estimated from the relationship between LD_{50} and bee weight. Nevertheless, weight is not the only driver of the LD_{50}, as demonstrated from another analysis which investigated generic 'intrinsic' sensitivity of various species (see Section 6.5.3.7 of the Supplementary document). Among those, *A. mellifera* was the most intrinsically sensitive, which gives confidence that the extrapolation factors from this species are likely protective, despite some uncertainty remains.

The whole analysis focussed on finding general patterns across substances. In some particular cases, a certain species may be more sensitive to a substance due to very specific toxicodynamic processes (Hayward et al., 2019). Nevertheless, at the current stage, this cannot be predicted or accounted for in the risk assessment, as the knowledge on this topic is still very limited.

Almost no information is available in the literature concerning the slope of dose-responses for bees other than honey bees. Nevertheless, there is no particular indication that the slope of the dose-response, which is mainly driven by toxicokinetic and toxicodynamic aspects, should vary significantly among bees. In consideration of this, as a pragmatic approach, it is suggested that the $slope_j$ obtained from tests carried out with honey bees is also used for assessing homologous effects to other bee groups (e.g. $slope_j$ obtained from chronic tests with honey bees may be used for determining the chronic dose-response of the other bee groups).

In case $slope_j$ is available from tests with (standard) bumble bees and/or solitary bees, this should be used as representative for their own group of bees.

6.6 Summary of the selection of hazard parameters for the risk assessment

To select the appropriate dose-response to be used for risk assessment for each group of bees and each risk case *j*, it is necessary to consider all the elements discussed above in relation to the two hazard parameters $LD_{50,j}$ and $slope_j$.

Since standard test protocols are available for honey bees covering the different risk cases, these are generally considered as starting points to derive the hazard parameters for other bee species as well, in consideration of the inter-species sensitivity. However, in some cases, other tests with bumble bees (*Bombus terrestis*) and solitary bees (*Osmia spp.*) standard species will also be available and should be used as a reference for the group of bees they belong to.

6.6.1 Hazard parameters for the risk assessment of honey bees

The first step to select the representative set of hazard parameters $LD_{50,j}$ and $slope_j$ for any test item (either an active substance or a PPP containing one active substance) is to consider all the available data. The procedure is summarised in Figure 5.

When multiple equivalent honey bee tests with the test item are available, priority should be given to those whose results allow the description of a full dose-response, i.e. those allowing the estimation of an uncensored LD_{50} and a slope, which should always be provided within the study reports included in the dossier (Section 6.2). When more of these tests are available, a geomean of both hazard parameters is suggested (Section 6.2).

However, if none of the available tests allow the description of a full dose-response, the one with the highest right-censored LD_{50} value should be considered (Section 6.2). In this case, a surrogate dose-response can still be obtained by applying the appropriate extrapolation factor to the censored LD_{50}, as described in Section 6.3. A worst-case default slope of 1.43 should be used in these cases.

It is noted that the selection of the hazard parameters for chronic risk assessment of honey bees should consider if the active substance shows time-reinforced toxicity (see Section 6.4). If this is the case, the life-span dose-response should substitute the 10-days chronic dose-response.

Finally, for identifying the most appropriate $LD_{50,j}$ for honey bees risk assessment of active substances, it is necessary to consider the difference in toxicity based on data for the representative PPP and the active substances (both expressed in terms of active substance). This step is not necessary if the representative PPP contains additional active substances. As explained in Section 6.1, if the $LD_{50,j}$ for the PPP is more than a factor of 3 below that of the active substance, the PPP is considered more toxic. Hence, the hazard parameters related to the PPP must be selected for the risk assessment of the active substance in the context of its inclusion/renewal. When one $LD_{50,j}$, between the active substance and the PPP is not a surrogate, this should be used for the risk assessment together with the related $slope_j$. When both active substance and PPP $LD_{50,j}$ are surrogate, additional case-by-case considerations must be made concerning the mortality actually observed in the tests and the tested doses. For example, if the top/limit dose caused no mortality for either the PPP or the active substance, a comparison is not meaningful. In such case, it would be appropriate to use the highest surrogate $LD_{50,j}$ for the risk assessment, together with a default $slope_j$ value of 1.43.

This comparison is normally performed on acute data; however, if the PPP is acutely more toxic, then also PPP chronic and larvae data should be provided and included in the comparison. In addition, if the PPP is more toxic, the risk assessment for the active substance should be based on the hazard parameters derived from tests with the representative PPP.

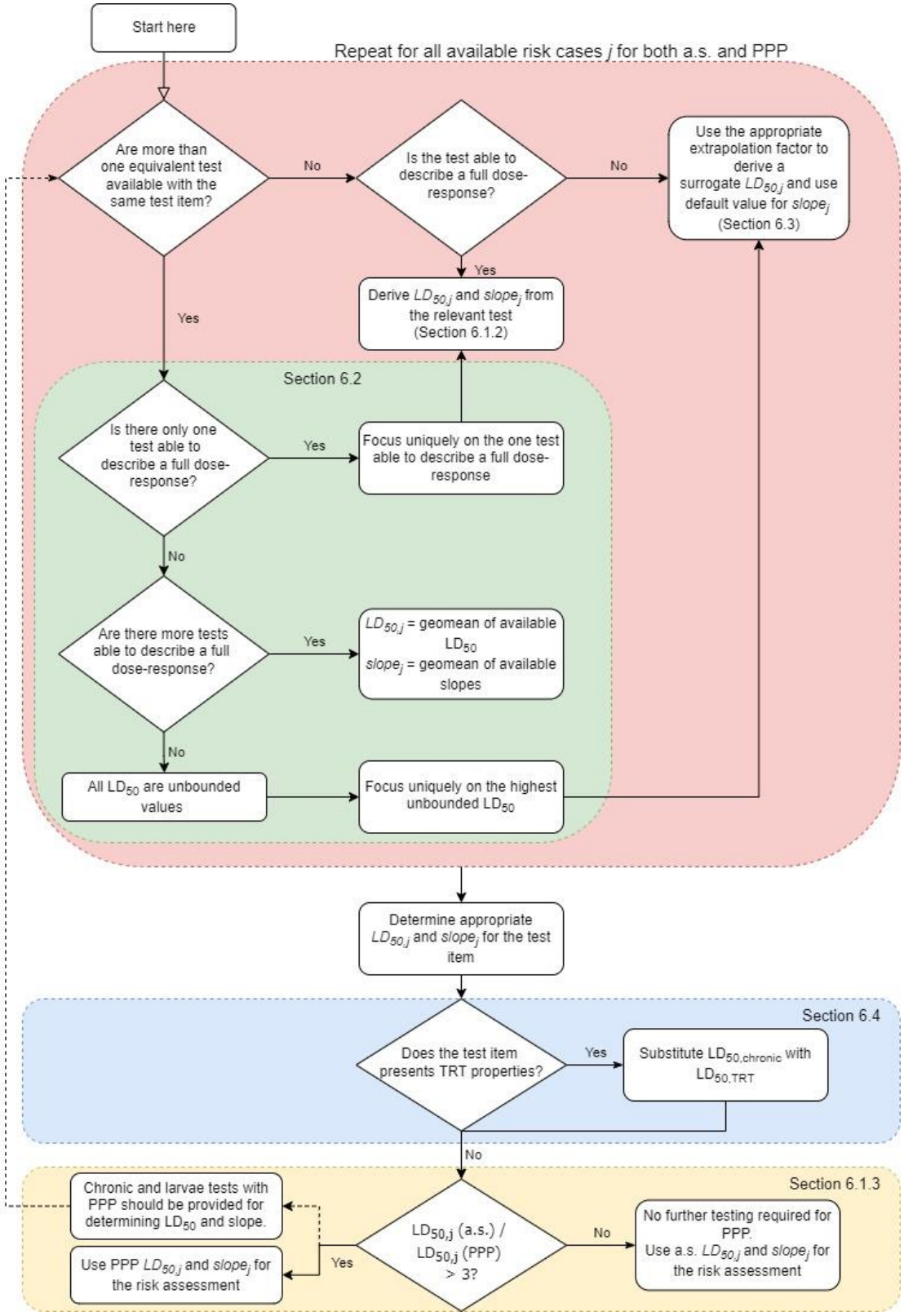


Figure 5: Flowchart illustrating the selection process of the hazard parameters for the risk assessment of honey bees. Note that the comparison between the active substance and the representative PPP is only meaningful

if the PPP does not contain additional active substances. If this is the case, the selection of the hazard parameters for the PPP should follow the indications reported for mixture risk assessment under Chapter 12.

6.6.2 Hazard parameters for the risk assessment of bumble bees

For bumble bees, OECD TG 246 and 247 acute tests may be provided in the dossier with both active substance and representative PPP. In addition, relevant literature data may be available. The treatment of the hazard parameters (i.e. $LD_{50,j}$ and $slope_j$) from any available test with the standard species (*Bombus terrestris* and, less frequently, *Bombus impatiens*) should follow the same process already summarised in Section 6.6.1 for honey bees (Figure 5), with the only exception that a TRT assessment is not applicable.

The selection process of the hazard parameters may include the combination of equivalent studies with the same species and the same test item, the definitions of surrogate dose-responses, and the comparison between the active substance and PPP data to identify the most appropriate $LD_{50,j}$ for risk assessment. All these steps are performed in a similar fashion to what was described for honey bees.

However, since the bumble bee group include many untested species (see Section 1), in order to cover the inter-species differences, toxicity extrapolation factors (*Tef*) as described in Section 6.5 should be applied to obtain the relevant *extrapolated* $LD_{50,j}$.

In all cases when bumble bee data are not available (e.g. likely always for chronic and larval effects) the *extrapolated* $LD_{50,j}$ should be obtained by applying the appropriate *Tef* to the $LD_{50,j}$ selected for honey bees.

If reliable slopes are available from tests with standard bumble bee species, these can be used as the representative $slope_j$ for the risk assessment. When these are missing, as it is likely to be case for chronic and larval studies, the $slope_j$ selected for honey bees should also be used for the risk assessment of bumble bees.

6.6.3 Hazard parameters for the risk assessment of solitary bees

For solitary bees, since standard tests are not yet available, the risk assessment should generally be based on hazard parameters previously selected for honey bees, with an *extrapolated* $LD_{50,j}$ obtained after applying the appropriate *Tef* to the honey bee $LD_{50,j}$, as explained in Section 6.5. The honey bee $slope_j$ can be used 'as is' for solitary bee risk assessment as well.

However, when studies based on publicly available test protocols or draft OECD TG e.g. on *O. rufa* are available (likely for acute exposure only), they can be used to derive the hazard parameters for the solitary bee risk assessment. When this is the case, the LD_{50} (or surrogate LD_{50}) from those studies could be used to obtain the *extrapolated* $LD_{50,j}$ after applying the appropriate *Tef*.

6.7 Options for refinement

As possible refinement of the hazard, the WG considered the use of two approaches when equivalent tests on different species are available for the same substance, namely:

- The geometric mean approach;
- The species sensitive distribution (SSD) approach.

The WG highlighted that the species commonly tested are limited to *A. mellifera*, *B. terrestris* and *Osmia* spp. Tests with these species are used for separate risk assessments for honey bee,

bumble bee and solitary bee, respectively. Any merging of data (either geometric mean or SSD) between these species is, therefore, not appropriate.

If studies are available with multiple species belonging to the same bee group, the geomean approach could in principle be used to combine LD_{50} values after having applied the appropriate *Tef* to each of them. However, no further specific guidance is given here. It is anticipated that this approach would not provide any significant change in the risk assessment, unless the generic influence of weight on the LD_{50} used in this guidance (see Section 6.5) is significantly altered for the specific substance being investigated.

Similarly, in principle, if at least five species for a group (e.g. five solitary bee species) are tested and their body weight is sufficiently representative of the weight range of all bees in a group (e.g. from ~5 to ~130 mg for solitary bees, from ~100 to ~500 mg for bumble bees) or it is proven that body weight is not a significant driver of the hazard definition, the SSD could be a valid approach. In such case, the HD_5 could be used as representative $LD_{50,j}$ for the group. However, the main practical issue preventing its use is the lack of standardised testing guidance for a sufficient number of species.

In the context of the present guidance document, it was not possible to investigate whether any relationship exists between the slope of the dose-response and biological traits of the bees. In principle, if, for a specific substance, a positive correlation between the slope and the bee weight is seen, such relationship should be used to predict the slope of the 5th percentile weight of solitary bees and bumble bees, following the same principle illustrated for the LD_{50}. However, the current complete lack of experience in this sense requires future guidance development. On the contrary, if such positive correlation is not recorded, the $slope_j$ parameter can in principle be obtained by calculating a geometric mean of the available slopes from the different species.

When taking together the lack of standardised test guidelines for many species and the general lack of knowledge on inter-species variability in the dose-response, the WG cannot currently recommend the use of either the geomean or the SSD approach for bees. Nevertheless, for the time being, hazard information for multiple species could be considered in a weight of evidence, acknowledging that increasing the current level of knowledge would certainly improve the accuracy of the risk assessment in future.

7 Lower tier risk assessment

In line with the principles of the tiered risk assessment approach, the aim of the proposed methodology is to translate the agreed SPG for the lower tier risk assessment, resulting in a conservative assessment which simultaneously identifies active substances of concern whilst excluding the substances of least concern from further evaluation.

The suggested approach leaves behind the paradigm of evaluating the single endpoints separately, and follows the rationale of combining a precise calculation of individual level effects with an extrapolation from individual to colony levels, and the aggregation of effects of a PPP on separate endpoints using the concept of response addition. This calculation method considers that in real life, PPPs can affect a honey bee colony by different routes of exposure (see Section 1.4). Here, instead of using empirical safety factors for assessing the risk of single exposure routes separately, this approach combines the predicted effects in a more mechanistic concept. This is much closer to the risk of PPPs in real life and it is in line with the defined SPG, where the attribute dimension is the colony/population (see Chapter 3). The predicted effects following the exposure to a PPP at colony levels can then be directly compared to the SPG for honey bees. This

means that the predicted effects will be compared to the agreed value of 10% of colony size reduction, which will be used a trigger.

Practically, the proposed procedure for a 'combined risk assessment' is composed of three consecutive steps:

1. Quantification of the effects at the individual level for each risk case based on standard laboratory ecotoxicological studies, and exposure estimates;
2. Extrapolation of the individual level effects to colony/population level effects for each risk case;
3. Combination of effects for all risk cases into a single predicted effect at the colony/population level.

These steps are described in the following section. The proposed methodology is suggested for use for all bee types, including honey bees, bumble bees and solitary bees, despite for bumble bees and solitary bees a threshold of acceptable effect was not defined.

7.1 Step-by-step explanation of the lower tier approach for honey bees

7.1.1 Step 1: Quantification of effects at individual levels

In this step, the hazard parameters $LD_{50,j}$ and $slope_j$ are used to define a dose-response relationship for each relevant bee type and life stage, which is used together with the relevant predicted exposure quantity (PEQ_j) to calculate the predicted effect at the individual level. This result describes the relationship between exposure to a PPP and the mortality and can be used to indicate the probability of death after exposure to a specified dose of the PPP, or the proportion of bees that would be expected to die after exposure to the same dose.

As explained in Chapter 6, the dose-response is assumed to be a 2-parameter log-logistic dose-response, described by the two hazard parameters $LD_{50,j}$ and $slope_j$. A non-linear dose-response relationship allows a much more accurate calculation of effects as compared to linear dose-response relationships (see Chapter 6). In the years since the publication of the previous guidance (EFSA, 2013), evidence has been collected that observed dose response data rather follow sigmoidal, not linear dose-response relationships (see Section 6.3 and Supplementary document Section 6.3).

Formally, the calculation of the predicted effect, in unit of percentage, is given by

$$PIE_j = 100 \cdot f(LD_{50,j}, slope_j, PEQ_j) = 100 \cdot \frac{1}{1 + \left(\frac{PEQ_j}{LD_{50,j}}\right)^{slope_j}}$$

where *PIE* is the predicted individual level effect following the application and exposure of a PPP, *j* refers to a risk case as assessed in an experimental test such as acute-contact, acute-dietary, chronic-dietary or repeated-dose-larvae; $LD_{50,j}$ and $slope_j$ are the relevant hazard parameters that parameterise a log-logistic dose-response function *f*, and PEQ_j is a realistic worst-case exposure estimate for the respective exposure assessment Tier, which could be screening, or Tier-1 or Tier-2. The exposure estimate is defined as ecotoxicologically relevant exposure quantity, here the uptake of pesticide by individual bee per time unit (see Figure 6). It is represented by the dose (PEQ_j) in the environment and the potency ($LD_{50,j}$) in the ecotoxicological experiment and they are given in the same unit, which is mass of a.s.·individual^{-1}

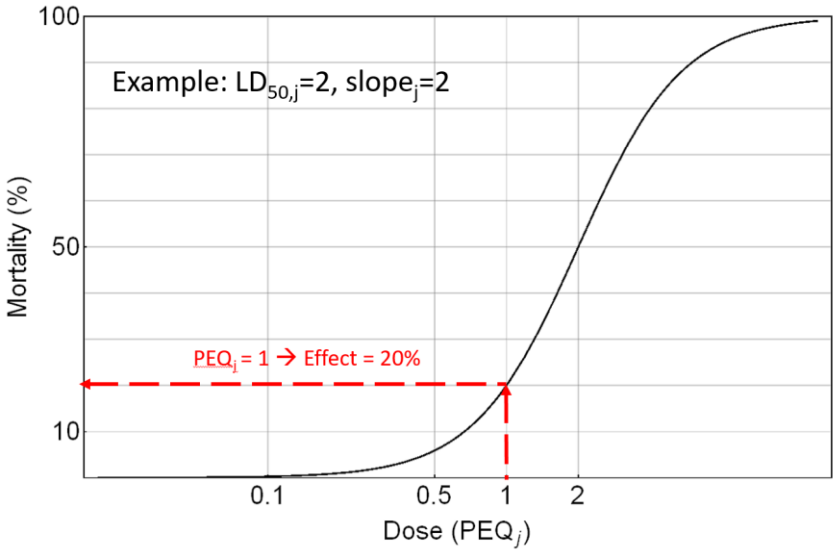


Figure 6: Graphical illustration of the proposed calculation for the effect on a specific endpoint using a non-linear dose-response curve. The resulting mortality (%) can be interpreted as probability of one individual to die on exposure to a certain dose, which can also be interpreted as a percentage of a cohort of individual bees to die after exposure to the identical dose.

The PEQ_j represents the most relevant exposure estimation of the various possible sources of ecotoxicological exposures as defined by a proper problem formulation (see Chapter 4, overview in Table 25). The calculation of PEQj values is described in Chapter 5, while the definition of the required hazard parameters, $LD50_j$ and $slope_j$, is in detail described in Chapter 6, including the determination of values for bumble bees and solitary bees, and the procedure if no proper dose-response relationship can be derived.

Table 25: Overview of exposure and the dose -response for the different life stages of honey bees

Life stage	Category	Exposure			Dose-response	
		Route	**Duration**	**Quantification and time scale**	**Potency**	**Slope**
Adult	Forager	Contact	Acute	From contact exp. model; pesticide mass sticking on the forager after a single application	Determined experimentally from the acute contact test. If the slope is not available, a worst-case value is used (see Chapter 6).	
Adult	Forager	Dietary (oral)	Acute	Worst-case between the two bee roles from dietary model; pesticide mass uptake per bee per day	Determined experimentally from the acute oral test. If the slope is not available, a worst-case value is used (see Chapter 6).	
Adult	In-hive (nurse)					
Adult	Forager	Dietary (oral)	Chronic	Worst-case between the two bee roles from dietary model; average daily pesticide mass uptake per bee during: 10 days (standard chronic assessment) - 27 days, i.e. the average lifespan of honey bee workers (for substance with TRT properties)	<u>Standard chronic assessment</u> Determined experimentally from the chronic oral test. If the slope is not available, a default worst-case value is used. <u>TRT assessment</u> Determined via extrapolation from the chronic oral test. The slope of the dose-response obtained at the end of the standard test is used for the lifespan dose-response as well.	
Adult	In-hive (nurse)					
Larvae	General worker	Dietary (oral)	Chronic (prolonged)	From dietary model; average daily pesticide mass uptake per larvae during 5 days	Determined experimentally from the larvae prolonged test/repeated exposure. If the slope is not available, a worst-case value is used (see Chapter 6).	

[a] When a substance has TRT properties, the risk should be evaluated for the entire honey bee lifespan for both the active (27 days) and the winter (182 days) period. Nevertheless, the winter scenario is a stand-alone assessment, which does not follow all of the steps illustrated in this Chapter. See Section 8.2.2 for more details.

7.1.2 Step 2: Extrapolation of the individual level effects to colony

In lower effect-tier assessment, toxicity endpoints investigated in lab studies are expressed at levels of individuals. However, the SPG defines the relevant ecological entities as the colony for honey bees and bumble bees, and the population for solitary bees. To make the lower tier risk assessment compliant with the SPG, effects need to be extrapolated from individual levels to higher levels of biological organization (i.e. colony or population). For honey bees, colonies include numerous individuals and a number of feedback mechanisms that are influenced by environmental conditions. Extrapolation from individual to colony levels for honey bees appears therefore reasonable to be considered, since not every individual level effect might immediately propagate equally to colony levels.

The extrapolation is here based on the general consideration of a worst-case exposure related to a certain risk case, here for larvae, foragers and in-hive bees.

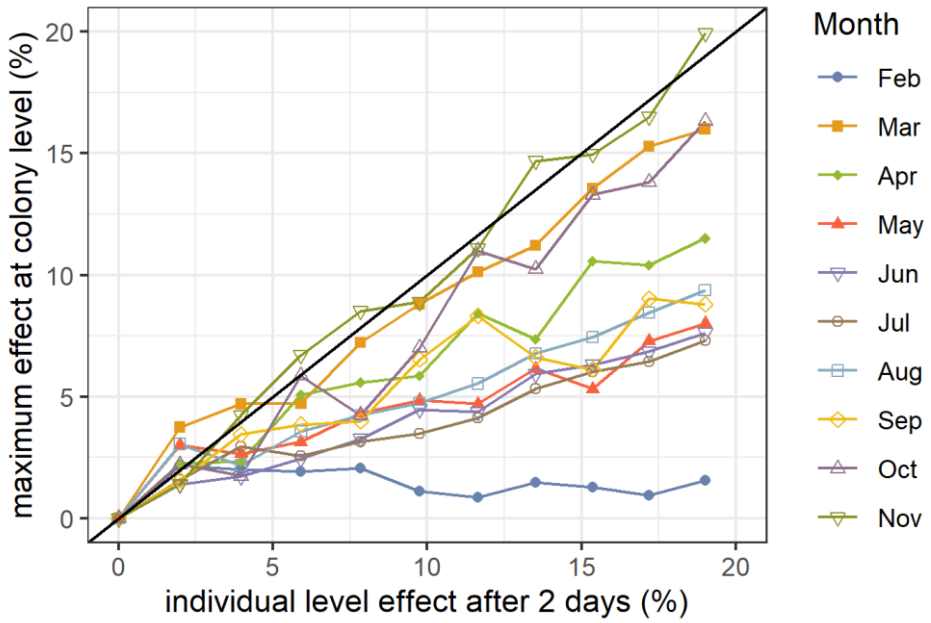


Figure 7: example BEEHAVE propagation: Example for the extrapolation of individual level mortality related to an assumed contact exposure to the colony level for the scenario D3 (EFSA et al., 2021). On the y-axis is the maximum observed effect at colony level after 2 days for an assumed contact exposure event in the middle of the respective 10 months in a year, indicated by single lines and symbols. The black solid line depicts the 1:1 line. It can be seen that for some months of the year, e.g. November, March and June, the effect at individual levels was more or less fully propagated to the colony level. The full set of simulation results from the BEEHAVE simulations are in the Supplementary document.

Contact exposure is assumed to occur only to foragers, so that extrapolation of effects from individual honey bee foragers to the colony level need to consider the different bee types within a honey bee colony and feedback mechanisms. The impact of contact exposure of forager bees on colony levels was therefore analysed by using the BEEHAVE model as the most suitable currently available choice (EFSA et al., 2021) under consideration of colony level feedback mechanisms and variable ecological conditions. In these simulations it was assumed that all foragers experience an increased mortality (without explicitly simulating any exposure), if they leave the hive at least for one foraging flight on a day with simulated additional mortality. In the vast majority of the simulated ecological conditions, the application of a certain level of increased forager mortality resulted in a correspondent lower effect at the colony level (i.e. in terms of strength). However, for some months of the year, the simulation results were close to the 1:1 line (Figure 7). Since in a Tier-1 method, an explicit consideration of application times for the extrapolation of individual to colony level impacts is not a suitable option, for foragers a 1:1 extrapolation is taken as default step for the extrapolation from individual to colony level.

Dietary exposure (acute, chronic, or larvae) is in this lower Tier method assumed to affect every bee in a colony via nectar and/or pollen, hence the 1:1 extrapolation from individual to colony level appears as only appropriate choice. These assumptions were checked for adult bees by simulations using the BEEHAVE model, where the 1:1 propagation was confirmed (see more details in the Supplementary document).

Therefore, in the lower Tier risk assessment, the extrapolation step assumes a conservative 1:1 propagation of individual to colony level effects for all experiments, i.e. using dietary and contact exposure, formally written as

$$PCE_j = PIE_j$$

where PCE_j is the predicted colony level effect obtained from risk case *j,* expressed as percentage.

7.1.3 Step 3: combination of effects at the colony

In this third step, effects predicted for single risk cases are combined, based on the rationale that under real world conditions effects produced by different routes of exposure and on different life stages are adding up at the colony level, which is the *ecological entity* defined for the SPG. This rationale is based on the assumption that different bees are exposed from acute or chronic, from dietary or contact. The addition of responses to the single risk cases as estimated in step 1, and extrapolated to the colony level at step 2, is mathematically expressed by:

$$PE_{SPG} = 100 \cdot (1 - \prod_{j=1}^{n}(1 - \frac{PCE_j}{100}))$$

where PE_{SPG} is the overall predicted effect at the colony level, in units of % of colony size reduction. This value is directly compared with the SPG i.e. ≤10% colony size reduction for honey bees,. The maximum effect is mathematically limited to 100%, independent of the number of considered endpoints. Neglecting the timing of single events in the response addition calculation is common for Tier-1 methods and a conservative assumption.

7.1.4 Quantification of the contribution of a risk case to the overall predicted effect

There might be cases, where the overall predicted effect at the colony level PE_{SPG} is dominated by a single risk case. Depending on whether or not a single risk case dominates the PE_{spg}, different options for refinement can be used in the higher tier risk assessment (see Chapter 10). In order to assess the potential dominance, the contribution of one risk case on the overall predicted effect needs to be quantified. This can't be done by simply building the ratio between a single risk case and the overall predicted effects, since the aggregation by response addition is not additive, but multiplicative. From the definition of the PE_{SPG}, a formula can be derived for the contribution of risk case *j* to the overall predicted effect

$$\Delta_j = \frac{Ln(1-PCE_j)}{Ln(1-PE_{SPG})}.$$

7.1.5 Sensitivity and impact analyses of the lower Tier method

The method for calculating lower tier risk estimates has introduced a fundamentally new paradigm for the evaluation of the risk to bees at the lower tiers. It follows a simple and transparent rationale, using accurate calculation rules, and avoids empirical assessment or safety factors. In order to check the appropriateness of this novel method, a sensitivity analysis and an impact assessment of the lower Tier risk assessment was performed (see Supplementary document, Section 7.1.5). In order to keep these analyses manageable, they were limited to the 'treated crop' scenario for attractive crops (at least for pollen) and to spray applications which would result in a non-negligible exposure via pollen and nectar for this scenario. While these limitations constrain considerably the scope of the exercise, they made the calculations manageable.

The lower Tier methodology as such is based on a fully parameterized dose-response curve per test, i.e. LD_{50} and slope values per each risk case. This will not always be achievable, but LD_{50} values will usually be available for honey bees, and default slope values can be used in case no solid dose-response function including a slope value can be derived. Next to the predicted PEQj values in the environment, this is basically all necessary information.

7.2 Implementation of the combined risk assessment in the tiered approach

As explained in Section 7.1 the quantification of individual effect (step 1 of the combined risk assessment) is driven by the PEQ_j from the exposure and by the $LD_{50,j}$, $slope_j$ for the hazard.

Since for the hazard there are no options for refinement (see Section 6.7), the lower tier risk assessment can be performed based on different exposure tiers (see Chapter 3). This would include a screening, tier 1 and tier 2 exposure tier for the dietary risk cases and a tier 1 and tier 2 for the contact risk case, as illustrated in the figure below.

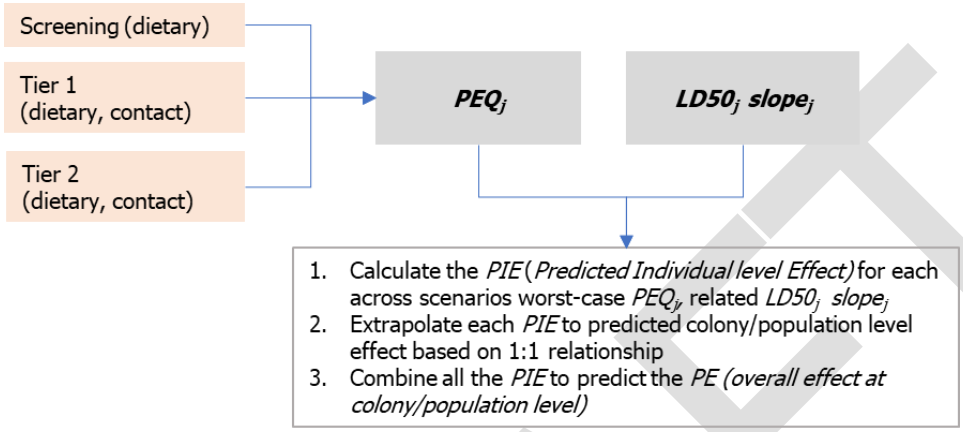


Figure 8: Combined risk assessment in relation to the exposure-tier

For the screening level risk assessment, the exposure estimation for the dietary route of exposure is done based on a simplified exposure model, as described in Section 5.6, resulting in more conservative exposure estimations compared to Tier 1. For the contact route of exposure, an additional simplified version of the exposure model is not available, and the same model and parameterisation as for the exposure estimation in Tier 1 (see Sections 5.1.1 and 5.2) is used. This exposure estimation is to be done for all relevant exposure scenarios (see also Section 4.3).

Based on the worst-case PEQ_i, the exposure scenario that dominates the exposure is selected (see also Section 4.3). For this scenario, the predicted individual level effect is calculated based on the PEQ_i and the related $LD_{50,j}$ and $slope_j$ values, for each of the risk cases (acute contact, acute dietary, chronic dietary and larvae) (see Section 7.1.1). The predicted individual level effects are then combined to determine the overall predicted effect at the colony level (see Section 7.1.3), which can be compared to the SPG. Note that if a substance presents time-reinforced toxicity, the risk assessment should be based at least upon Tier 1 exposure estimates (see below), i.e. a screening level risk assessment cannot be performed.

If a risk assessment indicates a high risk based on the screening exposure tier for dietary, a risk assessment with Tier 1 exposure estimations should be performed.

For the Tier 1 exposure tier, the dietary exposure estimation for the dietary route of exposure is done based on the full model described in Section 5.1.2. This more complex, but more realistic model assumptions result in a less conservative exposure estimation compared to the screening step. The parameters used in this model are described in Section 5.3. For the contact route of exposure, the model and parameterisation as described in Sections 5.1.1 and 5.2 is used, as in the screening level risk assessment. This exposure estimation is to be done for all relevant exposure scenarios (see also Section 4.3).

The next steps in the risk assessment (i.e. the effect evaluation) are performed in the same way as for the screening level assessment: For the dominant exposure scenario, the predicted individual level effect is calculated based on the PEQ_i, and the related $LD_{50,j}$ and $slope_j$ values, for each of the risk cases, which are then combined in the overall predicted effect at the colony level, which can be compared to the SPG (see Section 7.1).

If a substance is found to show time-reinforced toxicity, the predicted individual level effect for the chronic dietary risk case should be calculated differently: instead of the standard 10-day LDD_{50}, a lifespan LDD_{50} (covering a 27-day lifespan) should be used, together with a PEQ_i calculated for a 27-day exposure period (see also Section 8.2.1). An additional risk assessment for the winter period must be performed as well (see also Section 8.2.2).

If the risk assessment based on the -exposure-Tier 1 indicates a high risk (i.e. SPG not met), and it is not possible to mitigate the risk, a risk assessment based on exposure-Tier 2 is required. It is noted that if a proper exposure-Tier 2 assessment is not available, and the risk was not excluded at the lower tiers, the conclusion on the risk assessment will be drawn on the basis of those lower tiers.

At the Tier 2 exposure assessment, several of the parameters in both the contact and dietary exposure model can be refined. Refer to Sections 5.4 and 5.5 for details on the options for refinement and the need to generate further data. Using the refined parameter values, refined shortcut values can be calculated (see Section 5.5.4 for details), which in turn can be used in the model to calculate the Tier 2 PEQ_i.

The next steps in the risk assessment (i.e. the effect evaluation) are performed in the same way as for the screening level assessment (see Section 7.1), with again a different approach for substances that are identified to have time-reinforced toxicity properties (see Section 8.2.1 and 8.2.2).

If the risk assessment based on exposure-Tier 2 still indicates a high risk (i.e. SPG not met), a higher tier risk assessment has to be performed (see Chapter 10).

In summary, when a high risk at colony level is not excluded, any predicted individual level effect can be reiteratively refined according to the tiered approach. Higher tier effect assessment is needed when no options are available to refine the exposure estimation.

7.3 Implementation of the combined risk assessment approach for bumble bees and solitary bees

As mentioned in Section 3.1, the magnitude dimension of the SPG for bumble bees and solitary bees was discussed by risk managers on the basis on the consolidated evidence provided by EFSA (EFSA, 2022). MS risk managers agreed to select the option of 'undefined threshold' of acceptable effect indicated in the EFSA document (EFSA, 2022), and to require more frequently higher tier data allowing a better protection of these bee group, in the current absence of knowledge.

Based on an 'undefined threshold', a lower tier risk assessment scheme cannot be implemented since there are no values which would allow to interpret any quantitative lower tier outcome.

However, in this guidance, exposure estimation and hazard definition for bumble bees and solitary bees is possible, although these are characterised by considerable uncertainties due to the lack of specific data. This may have led to very conservative estimates in some cases.

Thus, in principle the combined approach described in Section 7.1 and its implementation in the tiered approach, can be applied also for these groups of bees to support the generation of higher tier studies. In other words, it is possible to estimate the predicted effects at colony/population level and to use this prediction as part of the problem formulation for generating higher tier studies, acknowledging that the magnitude dimension is defined qualitatively as 'no unacceptable effects', as indicated by general protection goal of the Regulation 1107/2009.

The implementation of the combined approach described in Section 7.1 to estimate the predicted effect at colony/population level may ensure that robust data are generated to better protect bumble bees and solitary bees. To implement such approach, applicant and risk assessors could refer to Chapters 5 and 6 for the exposure and hazard characterisation respectively, to apply the step 1 described in Section 7.1.1. Regarding the step 2 i.e. extrapolation of effect from individual to colony/population described in Section 7.1.2, the WG proposed to apply the same the 1:1 relationship relative to the propagation of effects from individual to colony/population since is considered conservative also for these bee groups. The combination of the effects (step 3) as described in Section 7.1.3 would allow also here to calculate the overall predicted effect on the basis of the addition of responses at colony/population level.

For illustrative purpose, the three steps described here for bumble bees and solitary bees were followed in the impact analysis performed on real uses as described in Section 7.1.5 and in more details in Section 7.1.5 of the Supplementary document.

7.3.1 Interpretation of the result

In lack of a risk assessment threshold, the overall predicted effect at lower tier must not be used to conclude on the 'acceptability of the risk'. However, it may be useful to trigger and to tailor further information which allow considering effects at higher level of biological organisation.

To this purpose, as reported in Duan et al. (2022), it should be considered that there are several ecological factors that could influence the vulnerability of bumble bees and solitary bees to PPP compared with honey bees (EFSA PPR Panel, 2012), tables 2.3 and 2.4). In general, biology and ecology of bumble bees and, especially of solitary bees, suggest a lower resilience and higher vulnerability to stressors relative to honey bees. Although it is unclear to what extent each ecological factor contributes to its vulnerability, it is important to highlight that this is a remaining source of uncertainty in the risk assessment that is difficult to quantify due to the lack of data. Therefore, a conservative approach to interpret the result, as also recommended by the opinion (EFSA PPR Panel, 2012), could be warranted. On the other hand, it is highlighted that specific aspects related to the biology and the interspecies sensitivity are already included in the exposure and hazard definition (see Chapters 5 and 6), although there are still uncertainties due to the lack of robust data.

8 Time-reinforced toxicity

A substance shows time-reinforced toxicity when its toxic effects from exposure to low doses for a long period of time are higher compared to effects from exposure to higher doses for a short period of time (i.e. its toxic effects are reinforced by exposure time). Note that this phenomenon was called 'accumulative toxicity' in the EFSA (2013).

The risk assessment scheme related to time-reinforced toxicity is shown in Figure 9. This scheme can be split into **two main parts**: a first part which focusses on the **hazard** and assesses

whether a substance shows time-reinforced toxicity; a second part which is the actual **risk assessment** for those substances that show time-reinforced toxicity.

The risk assessment part of the scheme is only relevant for substances that were found to show time-reinforced toxicity in the hazard part. When there are no indications that a substance shows time-reinforced toxicity, the standard Tier 1 chronic risk assessment as described in Chapter 7 is considered sufficient to address the risk from long-term exposure. However, if a substance is identified as showing time-reinforced toxicity, this is no longer the case, and the risk assessment below supersedes the standard chronic risk assessment.

Below, each part of the overall scheme is further described. For further background on the methods and the origin of the parameter values used, please refer to Annex G of the Supplementary document.

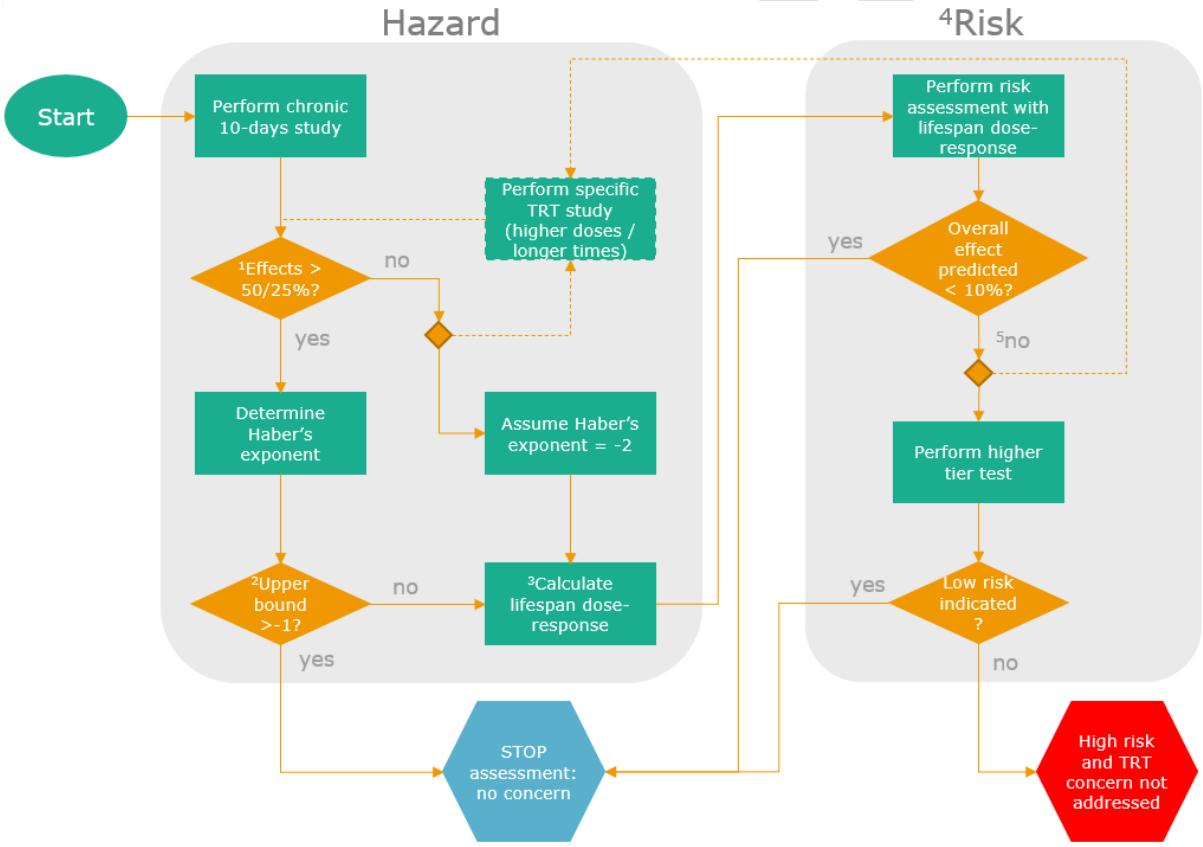


Figure 9: Flowchart of the scheme for the assessment whether a substance exhibits time-reinforced toxicity.

Notes:

[1] In the first diamond two questions are nested: are there at least 4 time points for which LDD_{50} can be calculated? If not, are there at least 4 time points for which LDD_{25} can be calculated?

[2] The second diamond includes the decision about quality criteria (R^2>0.85) and whether it is acceptable to use only LDD_{50} vs time or whether GUTS-SD should be fitted.

[3] Lifespan dose-response can be calculated with either using the Haber's exponent (worst-case) or GUTS-SD. In addition, it might be calculated for active period and winter scenario

[4] This part has to be duplicated for the active period and the winter (inactive) period

[5] When an effect >10% is predicted, either higher tier studies can be performed, or a specific TRT study can be executed. This second option is only applicable when this conclusion of >10% effect is reached on the basis of the worst-case assumption that Haber's exponent = -2.

8.1 Hazard assessment

The hazard assessment part of the flowchart shown in Figure 9 intends to solve the question whether or not a substance shows time-reinforced toxicity.

8.1.1 Estimating Haber's exponent

The starting point is the standard 10-day chronic honey bee toxicity study (OECD, 2017c). The first step is to check whether the data from the available 10-day chronic study allows to investigate the occurrence of time-reinforced toxicity using a C *vs* t log-log regression analysis. In order for such an assessment to be performed, a sufficient amount of mortality should occur in the study for a sufficient number of days. It should be possible to calculate an LDD_x value for at least 4 days. The LDD_x considered should preferably be the LDD_{50}. However, if an LDD_{50} cannot be calculated for at least 4 time points, the LDD_{25} could be considered instead.

If it is not possible to calculate the LDD_{50} or LDD_{25} for a minimum of 4 days, the substance shows too low mortality in the test, and it is not possible to estimate Haber's exponent based on this 10-day chronic study. In such cases, as shown in the flowchart (Figure 9), there are two options:

1. The first option is to perform **a new chronic study**. For the study design, there are again different options (which could also be combined): 1) use the standard 10-day chronic toxicity study test design (according to OECD 245), but use **higher treatment doses**; 2) Perform a study with a **longer duration** (i.e. extended beyond 10-days). This could be a study specifically tailored to address time-reinforced toxicity, in which a lower number of doses is tested compared to the standard 10-day test. In both cases, the aim is to increase the number of data points for the regression analysis. Refer to Annex G to the Supplementary document for some additional recommendations for the study design.

2. A second option is to **assume that the substance has a Haber's exponent of -2**, which can be considered as a worst-case screening step.

If an LDD_{50}/LDD_{25} can be calculated for at least four time points, the second step is to determine Haber's exponent (i.e. the slope of the regression line for the LDD_x *vs* time plot on a log-log scale). In a first instance, a regression analysis using the LDD_x values estimated for each time point separately (method B2-1 in Annex G to the Supplementary document) could be performed to determine Haber's exponent. If based on this method, the upper bound of the 95% confidence interval for Haber's exponent is below minus 1, it can be concluded that the substance shows time-reinforced toxicity. On the contrary, if the mean value of Haber's exponent is above minus 1, it can be concluded that the substances does not show time-reinforced toxicity. However, if the mean value is below minus 1 and there is overlap of the 95% confidence interval with minus 1, then in a second step also a calculation using GUTS (approach B2-2) should be performed. This calculation with GUTS-SD will reduce the uncertainty and will allow for a more robust conclusion. If based on the GUTS modelling, the upper bound of the 95% confidence interval is below minus 1, there is evidence that the substance shows time-reinforced toxicity. If this value is however equal to or higher than minus 1, there is no concern for time-reinforced toxicity.

It should be noted that an additional quality criterion for the regression analysis (for both method B2-1 and B2-2) was agreed: the **R^2 value of the estimated regression line should be > 0.85**. If this is not the case, the regression line is not considered reliable, and the data not suitable for assessing the occurrence of time-reinforced toxicity. Either a new study should then be performed, or Haber's exponent is assumed to be -2 in the further steps of the flowchart.

8.1.2 Calculating the lifespan dose-response (LDD_{50} and slope)

If the calculated upper bound of the 95% confidence interval of Haber's exponent is below -1, there is evidence that the substance shows time-reinforced toxicity. In that case, a risk assessment which covers the whole lifespan of a bee should be performed (see Section 8.2). In order to be able to perform such a risk assessment, the toxicity endpoint for a period of exposure that covers the whole lifespan should be known.

The WG agreed to consider two scenarios in the risk assessment: a scenario that covers the active period of the bees (i.e. summer), and an inactive winter scenario (for further details, see Section 8.2). A lifespan-LDD_{50} should therefore be calculated for both scenarios, using a lifespan of 27 and 182 days for summer and winter bees, respectively.

In principle, there are two ways in which the lifespan-LDD_{50} can be estimated. A **first method uses as a basis the log-log plot (LDD_{50} vs. time)** used for estimating the Haber's exponent. The derived linear regression is then used to derive the lifespan-LDD_{50}. It should be noted that in the risk assessment, all parameters of the dose-response relationship (i.e. not only the LDD_{50} but also the slope) are needed (see Chapter 6). Based on the LDD_{50} vs. time log-log plot, only the LDD_{50} for a given lifespan is calculated. An estimate of the corresponding slope is not possible. Therefore, the substance-specific slope for an exposure period of 10-days, as calculated from the available standard 10-day chronic toxicity study, should also be used for the longer exposure period considered in the lifespan risk assessment (which is a worst-case assumption).

A **second method** to calculate the lifespan-LDD_{50} **relies on predictions of the GUTS-SD model**. This approach has the advantage that for each time point a dose-response relationship is calculated, and that consequently both the LDD_{50} and slope are available.

In case the log-log plot is based on a LDD_{25} vs. time, the estimation of the LDD_{50} entails three steps:

1. Estimate the lifespan-LDD_{25} from the log-log plot.
2. Estimate the 10-days slope of the dose-response, if feasible, otherwise use a default value of 1.43 (see Section 6.5)

Calculate the lifespan LDD_{50} using the following formula

$$LDD_{50} = LDD_{25} \times \frac{1}{\left(-\frac{1}{(0.25-1)}-1\right)^{1/slope}}$$

When a default slope of 1.43 is chosen, the above is reduced to $LDD_{50} = LDD_{25} \times 2.16$.

Note that the lifespan LDD_{50} estimated using the log-log plot (LDD_{50} vs. time) will generally represents a worst-case compared to the lifespan LDD50 estimated using the GUTS-SD model, especially for the winter scenario (see Annex G to the Supplementary document, section 7.1.2).

8.2 Risk assessment based on TRT

For substances for which there is concern for time-reinforced toxicity following the first steps of the flowchart in Figure 9, the standard chronic risk assessment (see Chapter 7 for further details) might underestimate the risk from long-term exposure. Therefore, for such substances, a specific risk assessment, which covers the whole lifespan of a bee, should be performed. This specific risk assessment supersedes the standard chronic risk assessment.

As described in Annex G to the Supplementary document, the risk assessment as described below should only be performed for honey bees for the time being.

The endpoint to be used in this risk assessment is the lifespan-LDD_{50}. It was agreed to consider two scenarios in the risk assessment: one that covers the active period of the bees (i.e. a summer scenario), and a winter scenario. The input parameters to calculate the exposure and the different steps of the risk assessment are discussed below for both scenarios.

8.2.1 Risk assessment for the active period

For the lifespan risk assessment during the active period, it is assumed that a honey bee will live for 27 days. This value was derived from the data collected in the systematic review on bee background mortality (EFSA, 2020b). Based on this dataset, the median lifespan of bees during the active period is 27 days (90th perc= 41.7 days; 5th perc=17.5 days).

Given that the standard chronic risk assessment also focuses on bees during the active period, the same method for estimating the dietary exposure can be used in the lifespan risk assessment (see Chapter 5). The values for the different parameters (e.g. RUD and DT_{50} in pollen and nectar, daily food consumption, etc.), as used in the standard risk assessment, can also be used here. There are only two points specific for the TRT assessment:

- **The time window for calculating time-weighted average concentrations (called "*w*")**. Where the standard risk assessment considers a time window of 10 days (corresponding to the duration of the standard 10-day chronic oral toxicity study), a time window of 27 days (corresponding to the median lifespan, see above) is used in the lifespan risk assessment.
- **Pollen and nectar consumption**. During their entire lifespan, honey bee workers undergo changes in their diet in relation to the tasks they execute. Thus, for this specific case, a combination of subsequent diets was considered. Specifically, it was assumed that bees perform nursing activities for 10 days (pollen and nectar consumption), then 8 days of additional in-hive tasks (nectar consumption similar to the nursing phase, no pollen consumption), and 9 days of foraging activity (higher nectar consumption due to flying activities, no pollen consumption). See Section 5.3.4.5 of the Supplementary document for more details.

Taking into account the points above results in specific shortcut values, which are shown Appendix B of this guidance. These shortcut values are used to estimate the PEQ for the whole lifespan.

As in the standard chronic risk assessment, the predicted individual level effect is then estimated using the relevant PEQ and the lifespan dose-response (i.e. LDD_{50} and slope). This is combined with the other risk cases to estimate the overall predicted effect at the colony level (see Chapter 7).

8.2.2 Risk assessment for winter bees

During the winter period, honey bees will not forage for fresh pollen and nectar, but will feed on the stored food in the hive (i.e. honey). Therefore, whether honey bees are exposed to the pesticide depends on the presence of residues in the honey. Taking into account that beekeeping practices in Europe are highly variable, and that there are a lot of practices which can change the exposure over winter (e.g. dividing colonies), it is not possible to give a fully realistic estimate of the extent to which bees will feed on potentially contaminated honey. Therefore, for the lower tier risk assessment, it is assumed that all food during winter consists of contaminated honey.

For the winter bee scenario, a lifespan of six months (=182 days) was agreed by the WG.

As stated above, it is assumed for the lower tier risk assessment that food of the winter bees will consist of 100% contaminated honey. Based on this, the dietary exposure is estimated consistently with the exposure models presented in Chapter 5. Specific values for some parameters are defined irrespectively of the exposure model being used. These are:

- Sugar consumption from honey: in temperate regions, during winter, bees consume 8.8 mg of sugar/day to maintain the nest temperature at 5-8°C in the periphery and 15-20°C in the centre (Rortais et al., 2005)
- Sugar content in honey: as in EFSA (2013b), the water content of honey is assumed to be 18% (White, 1975). Therefore, the sugar content of honey would then be 82%.
- Dissipation rate in honey: Given that there is currently not data available on the DT_{50} of active substances in honey, it was agreed to use a worst-case value for the Tier 1 risk assessment (i.e. 1000 days), which results in no appreciable dissipation.

8.2.2.1 *Above-soil contamination*

The following equation is used to calculate the PEQ due to dietary exposure for the winter bee scenario in case of applications performed at BBCH stages >10 (spray applications and granules):

$$Rint = PEQ_{di} = AR\ SV$$

Where: *AR* = Application rate

SV = shortcut value for dietary exposure through honey

The shortcut values are calculated using the following equation:

$$SV_{winter,above} = \frac{1}{1000}\ CONC_h \frac{CMP_{su,w}}{S_h}$$

Where: $CMP_{su,w} = 8.8$ mg/day is the consumption of sugar in winter

$S_h = 0.82$ is the sugar content of honey

CONC$_h$ = Concentration in honey (mg/kg) = $RUD_h \frac{1-e^{-kw}}{kw}$, with $k = \ln(2)/DT_{50}$ and w = 182 days.

In lack of a suitable database for RUD values in honey (*RUD*$_h$), it was agreed to use the available database for RUDs in nectar as a surrogate in the Tier 1 exposure assessment. Based on data from residue trials in honey, available for six substances, this was demonstrated to be sufficiently worst-case (see Annex G of the Supplementary document). In case of granular applications at BBCH > 10, the RUDs for DW spraying are used as a surrogate, together with a dust formation factor of 0.1 and a safety factor of 3, consistently with the strategy applied for this application method in the derivation of the *Ef* described in Section 5.3.2.

It should be noted that in the other parts of the risk assessment, the SVs used for calculating the exposure are determined using a Monte Carlo simulation (see Section 5.3.4 for details). This allows to take into account the range and distribution of the values of the different parameters, and to derive an overall 90th percentile SV value. However, in the case of the winter bee scenario, there is only one parameter (RUD) for which a range and distribution is available. Therefore, using the 90th percentile value for the RUD would give a similar results for the SV than when a Monte Carlo simulation is performed. Given that the former is easier, the SVs for Tier 1 for the

winter bee scenario are obtained via a simple calculation according to the equation above, and by using the 90th percentile RUD values for nectar (12.2 and 15.2 mg/kg for DW and SW spraying, respectively). Using the values for the other parameters as mentioned above, this results in the following SVs shown in Table 26.

Table 26: Shortcut values for winter bees (above soil contamination)

Application	Shortcut value (µg/bee/day)
Downward spray	0.123
Sideward/upward spray	0.153
Granules	0.037

8.2.2.2 *Through soil contamination*

In lack of a suitable database for RUD values in honey, and consistently with what was agreed for the above-soil contamination model, it was decided to assume that the residue levels in nectar are an acceptable worst-case proxy for the residues in honey. As presented in Section 5.1.2.2, residues in nectar for contamination via soil can be estimated by using PEC_{pw}. Thus, the following equation is used to calculate the PEQ due to dietary exposure for the winter bee scenario in case of applications performed at BBCH stages <10 (seed treatment, spray applications, and granules):

$$Rint = PEQ_{di} = SV_{winter,soil} = \frac{1}{1000} \times PEC_{pw} \times \frac{CMP_{su,w}}{S_h} \times \frac{1 - e^{-kw}}{kw}$$

With $k = \ln(2)/DT_{50}$ and w = 182 days.

In the Tier 1 of the exposure assessment, consistently with what is reported in Section 5.3.15, the PEC_{pw} is assumed to be 1 mg/kg, for any application regime where the cumulative application rate is not higher than 4.5 kg/ha. In cases where the cumulative application rate is higher than 4.5 kg/ha, Tier 2 exposure estimation have to be conducted that requires a GAP specific PEC_{pw} calculation, as described in Section 5.5.15.

Note that all parameters used for Tier 1 exposure estimations are in this case fixed, therefore, the exposure estimation for winter scenario and through soil contamination model is 0.010 µg/bee/day.

As in the standard chronic risk assessment, the predicted individual level effect is then estimated using the relevant PEQ and the lifespan dose-response (i.e. LDD_{50} and slope). For the winter scenario, only continuous dietary exposure through honey consumption is considered. In addition, there are no larvae during winter. Therefore, the other risk cases (i.e. adult acute contact and dietary, larvae) are not relevant in this case. The winter scenario is thus a stand-alone assessment. The predicted individual level effect will correspond to the overall predicted effects at the colony level.

8.2.3 Refinement options

If the Tier 1 risk assessment for the active period and/or the winter bee scenario shows high risk (i.e. the SPG is not met), as next step, it is possible to refine the exposure estimation, according to the exposure tier and/or to perform a specific TRT study, as also described in Section 8.1.1. The latter would especially be useful for those substances for which initially it was not possible to calculate Haber's exponent, and it was therefore assumed that this Haber's exponent had a worst-case value of -2 for calculating the lifespan dose-response.

For the lifespan risk assessment for the **active period**, the exposure is estimated using the same dietary exposure method and parameters as in the standard risk assessment, with just minor differences (see Section 7.2.1). Therefore, all possible refinement options that could be used in the standard risk assessment could also be used here (see Sections 5.4 and 5.5).

For the lifespan risk assessment for the **winter bees**, the only two parameters in the dietary exposure model that could be refined using substance-specific data, are the residues in honey and the DT_{50} in honey. In case where the through soil model is applicable, a possible refinement for the residue level entails a better characterisation of the PEC_{pw} at the time of flowering, as explained in Section 5.5.15.

Data on the residues in honey could be obtained from available studies in the residue section of a pesticide dossier (e.g. submitted for setting an MRL in honey). However, experience with sampling honey for the bee risk assessment is relatively scarce, as during the active period, honey is not considered a good estimator of the exposure (see Annex B of the Guidance Document). During winter, some of the issues that honey presents during the active period (i.e. different level of maturation, concentrations and re-dilution) should be less relevant.

For both RUD and dissipation studies, it is important that samples combine different spots from different combs, in order to be representative of the whole storage. For dissipation studies, this strategy would also minimise the risk that samples are only representative for material (i.e. nectar) collected by the bees during a specific moment of the active period. Should this be the case, this would be a confounding factor for determining the dissipation kinetics.

For dissipation studies, it is also important to make sure that the active substance of concern is present in the capped honey cells in sufficient concentrations at the start of winter. Otherwise a suitable kinetic decline cannot be determined.

More specific guidance on the sampling strategy cannot be given for the time being, it is expected that further recommendations could be provided when more experience is gained.

The lifespan of a winter bee of 182 days is a very rough and conservative estimate. In theory, this value could be refined if more detailed data would be available. However, it should be noted that any refinement of the winter bee lifespan would have only a rather minor effect on the estimated lifespan-LDD_{50} (refer to Annex G to the Supplementary document for details). Therefore, this kind of (general) refinement is not considered very useful, unless the outcome of the risk assessment is borderline or it can be demonstrated that the length of the winter period is substantially less than three months, which is hardly the case for Europe.

8.2.4 Higher tier risk assessment

When a low risk could not be demonstrated in lower tier, the final step is to conduct higher tier effect field studies. The requirements for such studies are described in detail in Chapter 10. Generally, these requirements are the same for substances that show time-reinforced toxicity and those substances that do not. However, in case of a substance that shows time-reinforced toxicity (regardless of whether the risk assessment for the active period or the winter bee scenario fails at lower tier), a field study should be sufficiently long to ensure that potential effects following long-term exposure are addressed. In practice, this means that the study should not be started later than September and last until next spring, thus including overwintering and observing the honey bees for at least half a year.

9 Sublethal effects on honey bees in risk assessment

9.1 Overall strategy

The SPG for honey bees focusses on colony strength, and whilst the relationship between direct mortality and colony strength is relatively straightforward to establish (Section 3.1) the link between sublethal effects and colony strength is more tenuous. Despite this, concern for sublethal effects on wild and managed bee populations has been growing, primarily as a response to the broad range of harmful sublethal effects that have been reported across many bee species (see Annex K of the Supplementary document). Due to the growing concerns and specific calls for greater inclusion of sublethal effects to protect bee populations, the WG has decided to review the current provisions for assessing sublethal effects in the new guidance document.

Some major barriers to developing a comprehensive assessment of sublethal effects are their almost limitless diversity, the lack of standardization of measurement, and lack of a proven link to the SPG. To address these issues, the WG has decided to focus primarily on a subset of sublethal effects, in particular those that very obviously alter bee behaviour. The rationale for this is that large scale behavioural changes may interfere with important tasks, such as foraging, which would be expected to have a strong mechanistic link/relationship to colony strength. For example, a persistent sublethal effect that reduces the ability of honey bee workers to forage would reduce the amount of food available to a colony and the number of individuals in the hive which would eventually result in fewer bees and a smaller colony size that may eventually breach the SPG. Additionally, the WG believes that behavioural observations can be readily incorporated into standard laboratory tests, allowing applicants to identify potential concern for adverse effects on foraging behaviour. If a potential concern is raised, this can be further investigated in specific assays.

The WG have provided recommendations for a number of lower and higher tier tests; however, the pinnacle of the risk assessment process is still the field test. Ultimately, the goal of the risk assessment for bees is to show whether the SPG is met (see Chapter 3) and fully reliable and complete higher tier tests can directly demonstrate compliance with the SPG. Thus, at higher tiers if the SPG is met then even if sublethal effects occur they do not impair the colony strength sufficiently to be considered a concern. Therefore, if a reliable and complete higher tier risk assessment is available (see Chapter 10), there is no need to specifically consider sublethal effects and an assessment does not need to be carried out. In all other cases, a sublethal assessment focussed towards effects on foraging behaviour of honey bees is required. It is recommended to follow this strategy only once the standard lower tier risk assessment, based on mortality endpoints, shows low risk.

As internationally agreed test guidelines are mostly limited to honey bees (with the exception of acute toxicity testing for *Bombus* spp.), the WG agreed to limit the assessment for sublethal effects to honey bees for the time being. This will allow gathering experience with this new approach.

The strategy is shown in Figure 10 and explained below. Details on the assessment strategy are given in the Supplementary document Chapter 9.

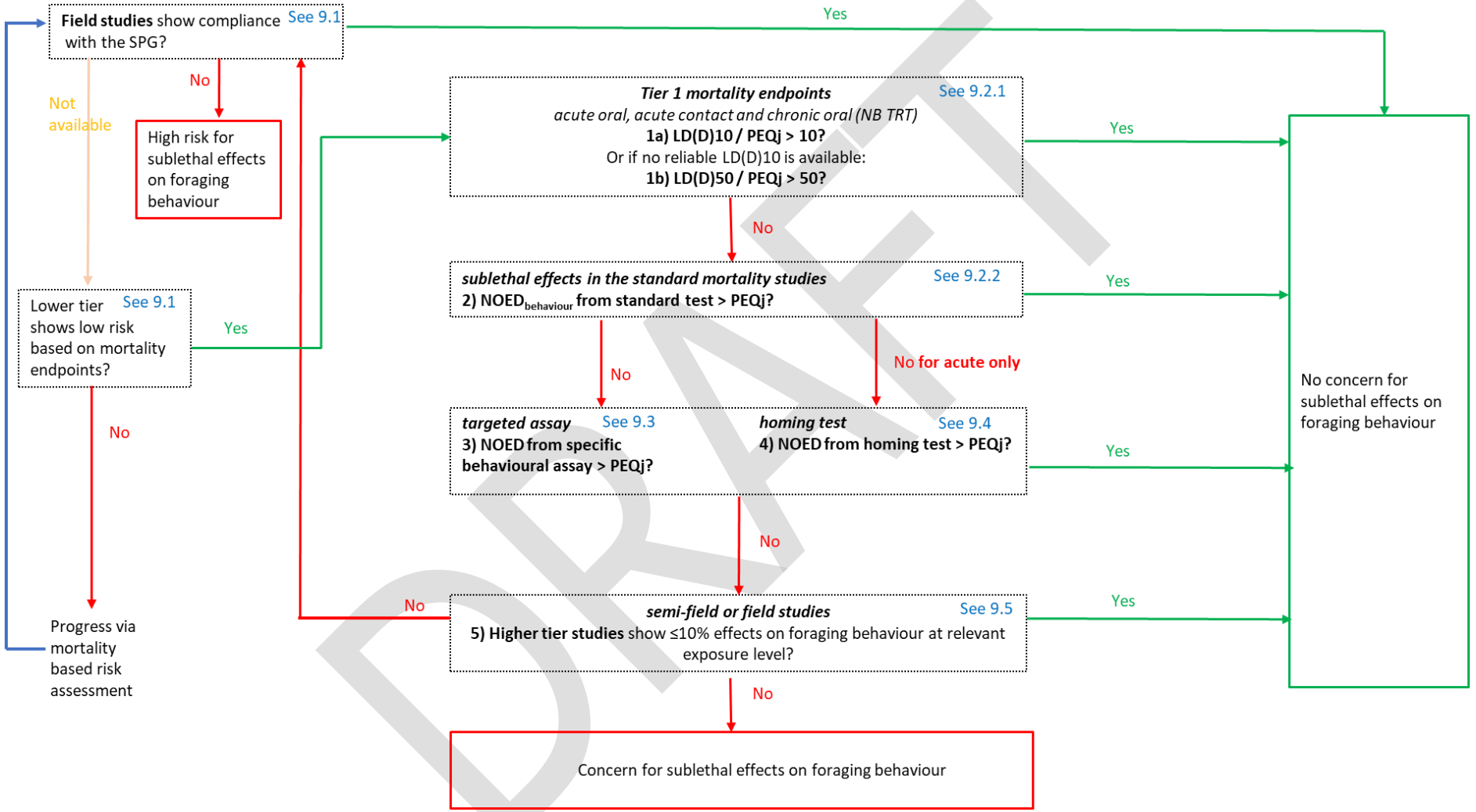


Figure 10: assessment strategy for sublethal effects

As our understanding of honeybee colony dynamics evolved it may be possible to use modelling to investigate the link between sublethal effects on foraging behaviour and colony strength, but this is currently not possible. Therefore, the WG acknowledges that in cases where a risk for sublethal effects on foraging behaviour is identified, the impact on colony strength is still unknown. The outcome of the sublethal risk assessment described here can therefore only be 'concern for sublethal effects indicated' or 'no concern for sublethal effects indicated'.

9.2 Strategy for triggering concern from lower tier information on honey bees

9.2.1 Toxicity/exposure ratio using mortality endpoints

The first step in the sublethal assessment is to estimate a "no concern level", i.e. a sufficiently low exposure (i.e. PEQ) which, if the bee is exposed to, no effect on foraging behaviour is expected. The "no concern level" can be calculated as the LD(D)10 divided by 10. If a reliable LD(D)10 cannot be calculated, the LD(D)50 divided by 50 can be used, which provides a more conservative no concern level.

The no concern level should be calculated separately for the acute contact, the acute oral and the chronic oral dose-response information and compared separately to the corresponding PEQ.

In step 1, a concern from sublethal effects in the lower tier is triggered by a PEQ which exceeds the no concern level calculated either by:

1a) Where a reliable L(D)D10 exists: a PEQj that is greater than the LD(D)10 / 10

1b) Where a reliable L(D)D10 does not exist: a PEQj that is greater than the LD(D)50 / 50

- PEQ_j (j indicates different PEQ values for the relevant risk cases) values are calculated according to Chapter 5. The most refined PEQ_j available, i.e. from the highest exposure assessment tier can be used.
- For limit tests with (virtually) no mortality, the following stands: LD(D)50 = LD(D)10 = NOED, meaning that the case 1a) is applicable.
- If a substance/product shows time-reinforced toxicity (TRT) and a time-to-effect study is available, the data from that time-to-effect study should be used. If no such test is available, the estimated lifespan LDD50 should be used (see Chapter 8 for more information). If a time-to-effect study is available for a substance that does not show TRT, the use of this test is not obligatory.

If the PEQ is lower than the "no concern level" (1a, 1b) for each of the three standard studies on dietary exposure of honey bees, i.e. acute contact, acute oral and chronic oral, then no concern for adverse effect for foraging behaviour can be concluded and no more consideration is needed. If the PEQ is higher than the "no concern level" a potential concern is identified and the risk assessor should consider the next step (step 2 in 9.2.2).

9.2.2 Using pattern of sublethal effects seen in the laboratory tests

In the second step, the no-concern level is calculated from proportion of abnormal behaviours observed during the test standard laboratory tests on honey bees (acute oral, acute contact and chronic oral) and the amount of food consumed by the bees. The WG propose using the regular observations required in OECD 213, 214 and 245 ((OECD, 2017c, OECD, 1998b, OECD, 1998a)) to determine if exposure to a PPP influences the behaviour of bees in laboratory experiments. As we cannot directly link any change in behaviour in a laboratory context to the SPG, then any statistically significant difference between the behaviour of the treatment and control groups

should be treated as indicative of a potentially important sublethal effect that requires further investigation.

Sections 9.2.2.1-9.2.2.3 explain how to derive a $NOED_{behaviour}$ from the standard laboratory tests.

2) If the $NOED_{behaviour}$ from the standard studies is higher than the PEQ_j, there is no concern indicated from sublethal effects on foraging behaviour.

If this is not the case, additional studies are needed, see step 3 in 9.3 and step 4 in 9.4.

9.2.2.1 *Statistical analysis of behavioural effects to derive no concern level ($NOED(D)_{behaviour}$)*

Tests such as OECD 245 can generate many observations which are not independent of each other, a phenomenon known as pseudoreplication. The presence of pseudoreplication violates the assumptions of many common statistical tests, namely that the data are independent, and therefore needs to be incorporated into the experimental design and analysis. The worked example in 9.2.2.3 describes a chronic exposure experiment with 1373 observations from 140 individuals spread over 14 cages repeatedly observed over 10 days. If the data is analysed using a chi square test, which can be used to test for differences in the proportions between groups, using all of the data would violate the assumption of independence; this would increase the chance of rejecting the null hypothesis, that there is no difference between the two groups, when in reality no difference exists (Type 1 error). Similarly, if we reduce the dataset to only keep independent observations, say by just analysing the observations from one randomly selected individual per cage on one specific day, this would so dramatically reduce the number of observations that it would be exceedingly unlikely we would observe a significant difference between the groups, even if one exists (Type 2 error).

Therefore, the WG recommend analysing the resulting data using statistical tests that incorporate any pseudoreplication (Zuur et al., 2009) that is present. There are well established classes of statistical tests that can incorporate pseudoreplication, in the example in 9.2.2.3, a generalised linear mixed effects model (GLMM) with a binomial error distribution was used to test for a difference in the proportion of abnormal behaviours observed between the lower concentrations of the PPP and a control group over 10 days.

The comparison should be performed for each test concentration below the median mortality endpoint (LD50, LDD50); at higher doses the mortality effects will take priority in the risk assessment. The treatment group consuming the highest concentration of the PPP where there is no difference between the treatment and control group can be taken as the no observable effect dose for behaviour, i.e. the 'no concern level'.

In summary, the procedure is as follows:

- The applicant should monitor the behaviour of the bees in each cage daily and report the total number of bees in the cage, the number behaving normally, number behaving abnormally (details on the behaviour to be observed are summarised in the Supplementary document Section 9.3.2).
- The data should be analysed using a mixed effect model comparing the treatment group to the control group.
- The model should incorporate the number of bees showing abnormal behaviour and the number showing normal behaviour as the response variable.
- The treatment (Control and the Treatment group), the observation day, and an interaction between the two should be included and the cage ID should be incorporated as a random effect.
- The error distribution should be binomial.

- There is a difference if either the main effect of treatment, or the treatment by day interaction is significant.
- To check that the random effects accurately reflect the blocking structure of the experiment, the number of observations should equal the total number of observations and the number of groups should equal the total number of cages.

The statistical analysis will show at which test dose level(s) the behaviour is significantly different from the control. The highest test dose level that shows no difference with the control, i.e. the $NOED_{behaviour}$, is taken as the no concern level.

9.2.2.2 *Statistical analysis of behavioural effects to derive no concern level for food consumption (NOED(D)behaviour,food)*

The applicant should also test to see if the PPP induces changes in food consumption using a similar analysis to the behavioural observations with the following modifications:

- The volume of food consumed in each 24 hour period should be recorded as the response variable
- The analysis should be carried out based on the amount of food consumed, standardised by the number of live bees in the cage (total volume of sucrose consumed divided by the number of live bees)
- The error distribution should be normal.
- There is a difference if either the main effect of treatment, or the treatment by day interaction is significant.

9.2.2.3 *Worked example of how to analyse the behavioural observations after exposure to a PPP.*

Data taken from a dossier study which has been anonymised. The question posed is whether significantly different behaviour is seen in any of the treatment groups below the LC50 value in a chronic oral adult honey bee test.

The number of abnormal observations, as a proportion of the total number of observations on live honeybees, was analysed using a GLMM. The model allowed a comparison between each treatment group and the control group. The control data showed no indication of any abnormal behaviour throughout the experiment (red line Figure 11) whilst all treatment groups showed a significant increase in the proportion of observations showing abnormal behaviour (**Table 27**). These results indicate that the no effect concentration is lower than the minimum concentration of 200 mg/kg, i.e. no $NOED_{behaviour}$ can be determined from this study.

Table 27: Model output from a binomial GLMM comparing the proportion of abnormal behaviours observed between treatment groups over 10 days. The three lines at the bottom (italics) indicate that there is a significant day by treatment interaction and that the slope of each treatment group differs from that of the control group.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-6.60694	0.51554	-12.816	<0.001
200	2.65088	0.98027	2.704	0.0069
500	4.69817	1.01504	4.629	0.0000
1250	-2.61346	1.04532	-2.5	0.0124
day	0.91271	0.07329	12.453	<0.001
200:day	*0.41963*	*0.14468*	*2.9*	*0.0037*
500:day	*-0.8016*	*0.14518*	*-5.522*	*0.0000*
1250:day	*0.41765*	*0.1445*	*2.89*	*0.0039*

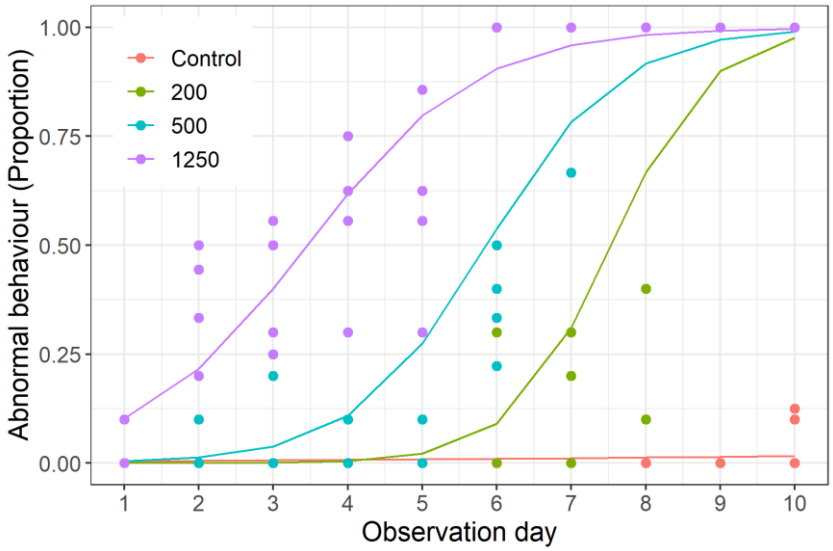


Figure 11: Each colour represents a different concentration (mg/kg). The individual circles represent the raw data and the lines represent the model predictions for the average response to pesticide exposure throughout the experiment.

9.3 Specific behavioural assays on honey bees

If a potential concern is raised based on the standard laboratory studies, targeted behavioural assays can be done. These studies should have similar study designs and statistical methods as the mortality tests but may use lower doses to include the no concern level. However, it has to be made sure that the predicted exposure of the exposure assessment tiers for the proposed GAP are covered in the test. In addition, they should implement the following two modifications that will improve the quality of the data on sublethal effects:

- Even with training, interpreting the behaviour of an animal can be subjective. In order to minimize any unconscious bias it is required that any behavioural results are generated blind (i.e. the observer does not know which treatment has been given to which group).
- The minimum number of replicates should be increased. The applicant has to demonstrate that the experiment is large enough to detect an effect size of at least 10% more observations of abnormal behaviour in the treatment group relative to the control group with an alpha of 0.05 and a power of 0.8.

In addition, the WG notes that using the OECD design, the behavioural observations can only be monitored as the proportion of individuals behaving abnormally at any point, making the unit of replication the cage. If individual bees can be marked, either by using paints or identification tags, then the behaviour of each individual can be recorded at each timepoint, potentially making the unit of replication the individual and increasing the statistical power of the test to detect an effect. This is however considered a potential future improvement, not implementable at present.

The modifications mentioned above (blind observer and increased number of replicates) can also be implemented directly in the mortality tests. The WG advises to consider these in future modifications of the OECD guidelines for acute and chronic adult bee toxicity. The $NOED_{behaviour}$ from the targeted assay should be compared to the relevant exposure level. This step supersedes the outcome of the previous steps.

3) If the $NOED_{behaviour}$ from the modified assays is higher than the PEQ_j of the exposure assessment tiers, there is no concern indicated from sublethal effects on foraging behaviour.

9.4 Homing flight study

The homing flight study is not required under the current data requirements, but an OECD guidance document has been published (OECD, 2021). The homing flight study can only be used as a refinement of a concern indicated from acute exposure.

4) If the $NOED_{homing\ test}$ is higher than the PEQ for acute exposure, there is no concern from acute sublethal effects on foraging behaviour.

9.5 Higher tier endpoints

If a fully reliable and complete higher tier risk assessment for honey bees shows compliance with the specific protection goal, apparently, even if sublethal effects occurred, they were not of such importance that colony or population size was impaired. However, a situation can be envisaged in which the available higher tier studies are not fully reliable to study effects in the longer term on colony strength but can still be used to study effects in the days directly after exposure. In such a case, assessment of effects on foraging behaviour can be useful and informative for the overall risk assessment.

In the absence of a set protection goal for sublethal effects, 10% effect is chosen as trigger, similar to the trigger for colony strength. In future, modelling (e.g. with ApisRAM) may be useful to investigate the link between these effects and colony size. An adverse effect on foraging behaviour is defined here as a 10% reduction (compared to the untreated control) of one or more of the following endpoints:

The amount of pollen collected per flight (in mass)

Ideally the amount of pollen is measured. However, it can also be estimated by proxy through the number of bees returning with pollen (see below). In future, the amount of pollen may be estimated through the size of the pollen pellet on the incoming bee as determined by analysis of video images. Current methods allow to count the number of bees returning with pollen in a set period (see below) and to weigh the amount of pollen collected in the pollen trap during that same period. By combining these, the calculation of the amount of pollen collected per flight can be estimated. For pollen traps, use the recommendations of Human et al. (2013) Delaplane et al. (2013). However, it is noted that it may be challenging to measure the amount of bees returning with pollen during the full period that the pollen trap is active.

The number of bees returning with pollen

This is considered as proxy for the amount of pollen entering the hive. It can be measured by visual observation of the hive entrance, by a human or automated observer.

For human observers, the recommendations of Delaplane et al. (2013) should be followed:

1. To control for between-colony variation due to time of the day, the investigator should: 1. limit observations to days and time of day with good flight conditions; 2. randomize the numeric order in which colonies are measured; 3. measure all colonies within a relatively narrow window of time, and; 4. limit colony observations to the same time window over successive days

2. Two observers sit to the side of a colony, each positioned well enough to the side to avoid obstructing the flight of the bees. Each observer has a hand-held counting device and one keeps time.
3. For one 15 min counting episode, each observer counts and records count foragers returning with, and without, corbicular pollen loads in order to derive proportion of foragers collecting pollen. Note that the length of the counting episode per hive may need to be adapted to reach a sufficient power of the study.
4. The mean of the two observers is derived and the data reported as total number of foragers returning and the proportion of foragers collecting pollen.

The observations should be done at the same time for paired control and treated fields.

Technology for automated observation is already available and will be further developed in the coming years. The WG was not in the position to give detailed guidance on the use of this technology, nevertheless the WG considered some of those example as promising tools for these observations. Therefore, it is expected that such tools will supersede the human observers method in the near future.

The duration of a foraging flight (in minutes/flight)

This parameter is considered useful in case the homing test indicates a concern. The parameters listed above also cover concern from the homing test, so if these are available, assessment of foraging flight duration is not obligatory.

The duration of a foraging flight can be assessed when individual foragers are marked. In order to assess a sufficient number of individuals, RFID (radio-frequency identification) technology is likely required. The recommendations of Scheiner et al. (2013) should be followed and the particular methods chosen from that Chapter should be justified.

The required endpoints can be measured in higher tier studies, both semi-field and field. The number of bees returning with pollen has to be assessed always, the amount of pollen is considered even more useful but likely not possible to measure yet, and the duration of a foraging flight is only obligatory if the homing test indicates a concern.

To measure the endpoints, the observations should be carried out at least 2 days prior to treatment (to demonstrate that the bees are acclimatized), and then just before and at several intervals after treatment (OEPP/EPPO, 2010). Observation timing will depend on the specific situation, but should include the days at which exposure in the higher tier effect study is expected to be highest. For a spray treatment, this is likely the day of application and the next three days afterwards. Observations should cover both the hours directly after application and the most active foraging hours of the day. For seed treatments and other pre-flowering applications, the peak of exposure is likely unknown a priori, the whole flowering period has to be covered by frequent observations, at least at every second day during the flowering period. As a general rule, the days and time of the day with good flight conditions should be selected for these observations. Effect calculations should be presented and evaluated per day.

The statistical analysis should be conducted according to the general approach for higher-tier studies (see Annex C). The applicant should apply a one-sided equivalence test ($\alpha = 0.2$) for each endpoint, with an equivalence limit corresponding to a 10% reduction in the treated test compared to the control, to prove that there are no adverse effects on foraging behaviour. Based on the current data it is not possible to provide indications on the number of replicates required to reach an adequate statistical power of the test.

Two additional endpoints were considered but not included as requirement. The density of foragers on the crop (bees/m2) was considered informative in the past for showing that exposure took place in a test. However, for the purpose of the sublethal assessment it is not considered useful, since the data are usually scarce and it is difficult to statistically compare controls with treatments. The amount of nectar collected (µL/flight) is considered a very useful endpoint. However, it is not included for the moment, because of the technical challenges encountered with the methods currently available.

In conclusion, step 5 of the assessment strategy is:

5) The foraging behaviour endpoints have to be investigated in higher tier studies at an exposure level that corresponds to the proposed GAP and should be sufficiently high, i.e. it should represent at least a 90th percentile exposure as predicted by one of the exposure assessment tiers. For more details, see Chapter 10. **If no effect >10% was seen at this exposure level, there is no concern from sublethal effects on foraging behaviour.**

If still a concern is indicated after this step, the only possible refinement is field studies that show compliance with the SPG (see Section 9.1 above).

10 Higher tier risk assessment

10.1 Introduction

In the higher tier of the risk assessment, the effects on bees are studied under more realistic conditions than in the lower tiers. The scope of the higher tier is to evaluate the risk under realistic use conditions, and finally check if the agreed SPG is met.

In this guidance document, three types of higher tier effect studies are considered, field studies, semi-field studies and colony feeder studies. Colony feeder study exists only for eusocial bees that form colonies. In the sections below, the selection of the study type with a brief description of the methodology and the main considerations, the aim and the endpoints will be explained. Internationally agreed and widely used guidelines (e.g. EPPO 170/2010, Lückmann and Schmitzer, 2019), – as far as it was possible - are largely followed. Detailed guidance is given in Annex C.

The aim of the higher tier effect studies is to measure the magnitude of effects on the study endpoint(s) after a field realistic worst case exposure to the PPP use under evaluation. Since the relevant endpoints under field realistic exposure conditions can be best studied in field conditions, the most suitable option is conducting a field study. However, under specific circumstances (see section below and Figure 2 in Chapter 3), the other study types could be considered. In all cases, the outcome of certain measurements (i.e. endpoints) from the exposed group(s) is compared with control group(s). All conditions of the exposed and control groups should be sufficiently similar with the only important difference being the PPP treatment itself.

The statistical analysis is based on the test of equivalence. This marks a substantial departure from the most common approach followed for effect studies. Indeed, the equivalence approach, is based on the assumption that the PPP causes adverse effects, since risk was not excluded at lower tier. In a standard approach, the starting hypothesis is that the PPP has no effect – that there is no risk. In the standard approach, evidence can only indicate - with a pre-established level of confidence - the presence of an effect of an undefined magnitude. When this is not the case, the starting hypothesis is not disproven, and the results are often interpreted as a complete

lack of effects and thus as a 'proof of safety'. However, such conclusion is not supported with the same level of confidence. In fact, the level of confidence for such conclusion can hardly be set a-priori, and even the evaluation a-posteriori presents challenges and pitfalls. A test of equivalence reverses the points of view, placing the burden of proof on the 'safer' hypothesis. The starting assumption is that the effect is larger than the SPG – that there is a high risk. The experimental evidence is used to prove the opposite, that the effect (if any) is smaller than the SPG – that the risk is low. If there is not enough evidence, the conclusion is that the risk is high, which is considered an appropriately conservative outcome. As a direct consequence, there is no need for risk assessors to recommend minimum levels of replications or sample size in the study. In summary, with the equivalence test it is possible to statistically support that there is low risk, which cannot be done with the standard approach. The statistical methods are described in Section 2 of Annex C –. Further details on the statistical requirements are included in the detailed description of each study type included in the same annex.

10.2 Higher tier studies for honey bees

10.2.1 Honey bee field study

Circumstances when the study type is useful

The study type is recommended when the lower tier risk assessment indicate no clear dominance of an exposure route or no clear dominance of a risk case and the impact on colony size (> SPG) is indicated to arise from 1) a combination of different routes of exposure, or 2) a combination of effects on the adults and larva, 3) or solely dietary risk on adults.

Methodology to be followed

Colonies with free flying bees are studied in open field conditions at field sites in different agricultural landscapes. The test colonies are located at the edge of the treated or the control fields. In order to confirm sufficient exposure, residues entering the hive have to be measured (see Sections 10.5 and 10.6.8)). The endpoint(s) has to be studied for all colonies in the same way.

Endpoint to be studied

The mandatory endpoint that should be studied, and which is the subject of the statistical comparison is the colony size, characterized by the number of adult bees in a colony.

Further definitions and detailed descriptions of the requirements are included in Sections 4.1 and 4.2 of Annex C –.

10.2.2 Honey bee semi-field study

Circumstances when the study type is useful

A) The study type is recommended when lower tier risk assessment indicates that the predicted effects on colony size (> SPG) are clearly dominated by the risk arising from the contact exposure.

B) The study type is recommended when, in the lower tier risk assessment, a concern for sublethal effects was flagged.

Methodology to be followed

Bee colonies enclosed in large cages are studied. The cages are in field conditions and include a flowering crop, which is either treated with the test PPP or treated with water (control) and treated with a substance toxic to bees (positive control). In order to prove sufficient exposure

(see also Sections 10.5 and 10.6.8), the foraging activity has to be observed. The endpoint(s) has to be studied for all colonies in the same way.

Endpoint to be studied

A) The mandatory endpoint that should be studied, and which is the subject of a statistical assessment is the forager mortality.

B) The endpoint to be studied, and which is the subject of a statistical assessment is the foraging behavior.

Further definitions and detailed descriptions of the requirements are included in Section 4.1 and 4.3 of Annex C –.

10.2.3 Honey bee colony feeder

Circumstances when the study type is useful

The study type is recommended when lower tier risk assessment indicate that the predicted effects on colony size (> SPG) is clearly dominated by the risk arising from the effect on larvae.

Methodology to be followed

Free flying bee colonies are studied in open field condition, but in a location with limited food resources. The test colonies are fed with sugar solution spiked with the test PPP or with sugar solution without the PPP (control) and treated with a substance toxic to bees (positive control). The exposure is controlled by the concentration(s) of the offered solution (see also Sections 10.5 and 10.6.8). The endpoint(s) (see below) has to be studied for all colonies in the same way.

Endpoint to be studied

The mandatory endpoint that should be studied, and which is the subject of the statistical comparison is the number of covered brood cells.

Further definitions and detailed descriptions of the requirements are included in section 4.1 and 4.4 of Annex C –.

10.3 Higher tier studies for bumble bees

Field studies and semi-field studies for bumble bees aim at investigating the same endpoints. However, those study types are performed in different conditions. The field test is the most realistic assessment of a PPP. Nevertheless, pending on the problem formulation and the PPP under evaluation, semi-field test might be conducted.

10.3.1 Bumble bee field study

Circumstances when the study type is useful

The study type is recommended when high risk with the lower tier risk assessment cannot be excluded and the lower tier risk assessment indicate no clear dominance of any of the exposure routes.

Methodology to be followed

Colonies with free flying bee are studied in open field conditions at field sites in different agricultural landscapes. The test colonies are located at the edge of the treated or the control fields. In order to prove sufficient exposure, residues entering the hive has to be measured (see Section 10.5, below). The endpoint(s) has to be studied for all colonies in the same way.

Endpoint to be studied

The mandatory endpoints which should be reported and analysed will be 1) the colony weight through time (weeks 0-8), which represents the colony strength, and 2) the number of queen producing cocoons present in the final census, which represents the colony reproductive output.

Further definitions and detailed descriptions of the requirements are included in Chapter 5 of Annex C –.

10.3.2 Bumble bee semi-field study

Circumstances when the study type is useful

The study type is recommended when high risk with the lower tier risk assessment cannot be excluded and the lower tier risk assessment indicate no clear dominance of any of the exposure routes.

Methodology to be followed

Bee colonies enclosed in large cages are studied. The cages are in field conditions and include a flowering crop, which is either treated with the test PPP or treated with water (control) and treated with a substance toxic to bees (positive control). In order to prove sufficient exposure, the flight activity has to be observed (see also Section 10.5, below). The endpoint(s) has to be studied for all colonies in the same way.

Endpoint to be studied

The mandatory endpoints which should be reported and analysed will be the colony weight through time (weeks 0-8) which represents the colony strength and the number of produced queen cocoons present in the final census which represents the colony reproductive output.

Further definitions and detailed descriptions of the requirements are included in Chapter 5 of Annex C –.

10.3.3 Bumble bee colony feeder study

Circumstances when the study type is useful

The study type is recommended when high risk with the lower tier risk assessment cannot be excluded and the lower tier risk assessment indicate clear dominance of the dietary exposure route.

Methodology to be followed

Free flying bee colonies are studied in open field condition. The test colonies are fed with sugar solution spiked with the test PPP or with sugar solution without the PPP (control) and treated with a substance toxic to bees (positive control). The exposure is controlled by the concentration(s) of the offered solution (see also Section 10.5, below). The endpoint(s) (see below) has to be studied for all colonies in the same way.

Endpoint to be studied

The mandatory endpoints which should be reported and analysed will be the colony weight through time (weeks 0-8) which represents the colony strength and the number of produced queen cocoons present in the final census which represents the colony reproductive output.

Further definitions and detailed descriptions of the requirements are included in Chapter 5 of Annex C – of this Guidance Document.

10.4 Higher tier studies for solitary bees

Also field studies and semi-filed studies for solitary bees aim at investigating the same endpoints. However, those study types are performed in different conditions. The field test is the most realistic assessment of a PPP. Nevertheless, pending on the problem formulation and the PPP under evaluation, semi-field test might be conducted.

It is also noted that the test species for which guidance can be provided, based on the current available knowledge are *O. bicornis* but *O. cornuta* or other solitary bee species could be considered. Further considerations to address the uncertainties around the extrapolation of the results to the numerous EU solitary bee species are strongly encouraged.

10.4.1 Solitary bee field study

Circumstances when the study type is useful

The study type is recommended when high risk with the lower tier risk assessment cannot be excluded.

Methodology to be followed

Free flying bees at nesting units (bee populations) are studied in open field conditions at field sites in different agricultural landscapes. The test bees are located at the edge of the treated or control fields. To evidence sufficient exposure, residues entering the nest must be measured (see Section 10.5, below). The endpoints (see below) have to be studied for all populations in the same way. The primarily test species for which guidance is provided is *O. bicornis* but *O. cornuta* or other solitary bee species could be considered.

Endpoint to be studied

The recommended endpoints that should be studied, and which is the subject of the statistical comparison are:

- Nesting activity as the proportion of nesting females in relation to the emerged number of females in the starting population
- Reproduction in the form of:
  - Number of brood cells
  - Number of cocoons
  - Number of emerged females and males in the next generation
- Population growth rate as the multiplicative product of nesting activity, number of offspring and proportion daughters

It should be noted that the recommended endpoints above are relevant for *Osmia* spp. In case other test species are used, the endpoints might need to be adapted to the biology/ecology of the test species; nevertheless they should be related to the next generation population size.

Further definitions and detailed descriptions of the requirements are included in Chapter 6 of Annex C –.

10.4.2 Solitary bee semi-field study

Circumstances when the study type is useful

The study type is recommended when high risk with the lower tier risk assessment cannot be excluded.

Methodology to be followed

Bees (bee populations) enclosed in large cages are studied. The cages are in field conditions and include a flowering crop, which is either treated with the test PPP, treated with water (negative control) and treated with a substance toxic to bees (positive control). To evidence sufficient exposure, the flight activity must be observed (see also Section 10.5, below). The endpoints (see below) have to be studied for all populations in the same way. The primarily test species for which guidance is provided is *O. bicornis* but *O. cornuta* or other solitary bee species could be considered.

Endpoint to be studied

The recommended endpoints that should be studied, and which is the subject of the statistical comparison are:

- Nesting activity as the proportion of nesting females in relation to the emerged number of females in the starting population
- Reproduction in the form of:
  - Number of brood cells
  - Number of cocoons
  - Number of emerged females and males in the next generation
- Population growth rate as the multiplicative product of nesting activity, number of offspring and proportion daughters

It should be noted that the recommended endpoints above are relevant for *Osmia* spp. In case other test species are used, the endpoints might need to be adapted to the biology/ecology of the test species; nevertheless they should be related to the next generation population size.

Further definitions and detailed descriptions of the requirements are included in Chapter 6 of Annex C –.

10.5 Exposure in higher tier effect studies *vs* ExAG

As explained above, higher tier studies, depending on the study type and on the organisms tested, provide a range of effect endpoints. In such studies, effects need to be investigated at an exposure level in line with the ExAG, i.e. 90th percentile worst-case exposure for the compound under evaluation. Therefore, appropriate exposure regimes and levels, defined EED_j in Chapter 3, have to be ensured in the study and expressed in the same unit as the PEQ_j. This means that, in addition to the biological observations, it is necessary to verify consistently the exposure levels e.g. via measurement of residues in pollen and nectar, and use these measurements to estimate the EED_j in the specific study. The EED_j is compared with the related PEQ_j in order to assess the plausibility of biological observations. This is further explained in Section 10.6.8. The comparison should be carried out with PEQ_j based on independent measured residue trials (e.g. Tier 2), but if not available, the PEQ_j from lower tier exposure assessment will be used. This is illustrated in Figure 12.

For the estimation of the EED_j the exposure models described in Chapter 5 has to be used, with the specific parameters measured in the effect studies. Default parameters can be retained while any parameters that is measured in the effect studies, in line with the recommendations for exposure-tier 2 refinements, can be applied.

It is noted, that a specific estimation of the EED is not necessary for semi-filed studies, since the foraging activity (number of honey bees foraging in the crop) or the flight activity at the colony/nest entrance (the total number of bumble bees or solitary bees flying in and out) is considered as proxy. These observations are used in order to show that the bees are actively foraging on the crop, and are therefore exposed to the PPP. In this case the EED correspond to the full amount of PPP applied in the study, which should be in line with the GAP under evaluation.

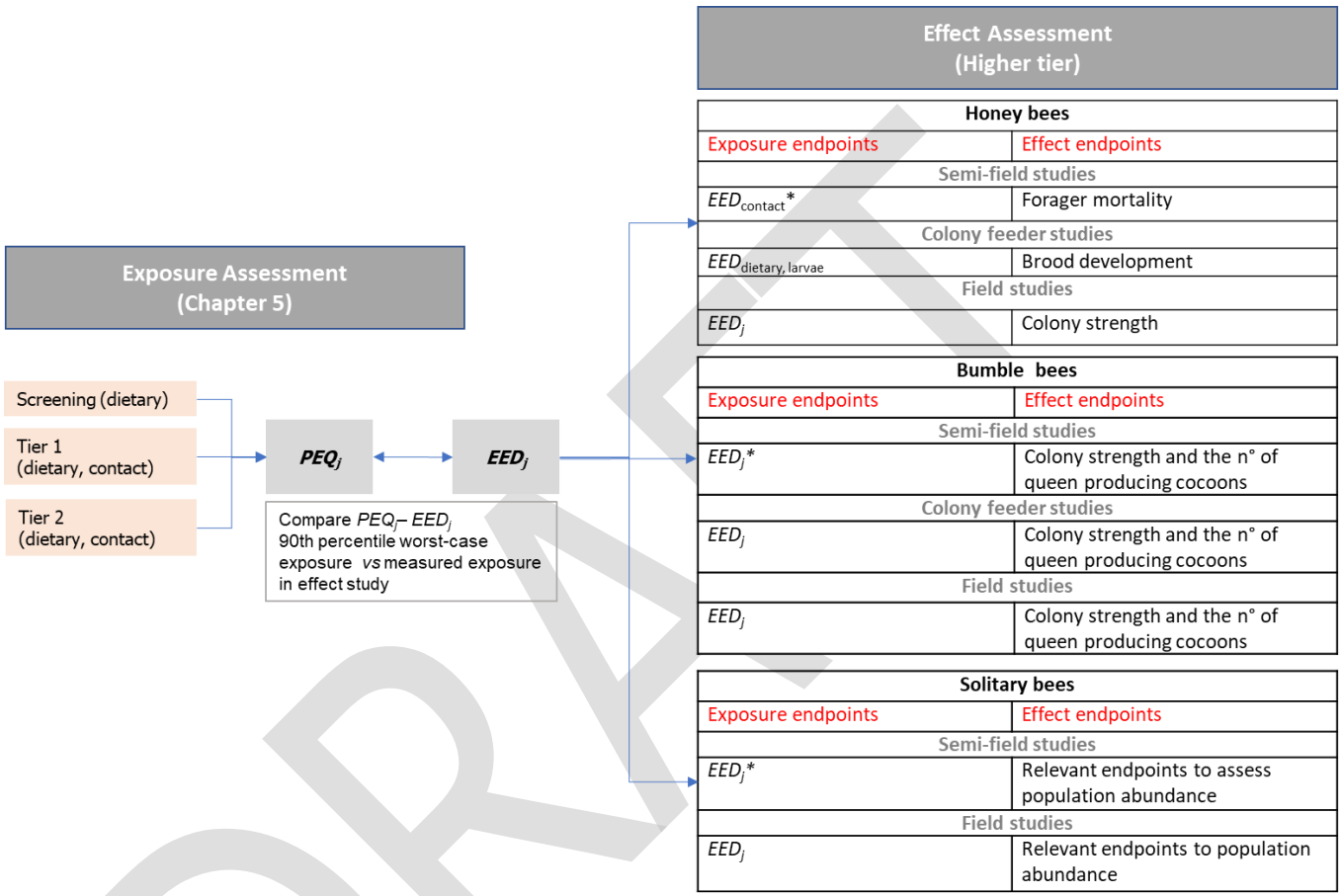


Figure 12: Exposure regime in higher tier effect studies should be estimated and compared with the PEQ in order to assess the plausibility of the biological observations in the studies. The suffix 'j' indicates the 4 risk cases (acute contact, acute, dietary, chronic, larvae). *The EEDcontact in semi-field studies, does not require to be estimated since the exposure level in the study is confirmed by the flight activity as described in Annex C of this guidance document

10.6 Weight of Evidence and uncertainty analysis

In this guidance document a structured weight of evidence (WoE) approach is recommended for evaluating the outcome of the effect higher tier studies.

The following section is largely based on the recommendations included in the EFSA Guidance on the use of the weight of evidence approach in scientific assessments (EFSA Scientific Committee et al., 2017). While not aiming at being highly prescriptive, some additional guidance is provided to make the generic principles of EFSA Scientific Committee et al. (2017) more specific to the pesticide risk assessment to bees.

10.6.1 Definitions and structure of the WoE

The building blocks of the weight of evidence process are generally referred to as '**lines of evidence**'. EFSA Scientific Committee et al. (2017) defines the line of evidence as 'a set of evidence of similar type'. In the context of the bee risk assessment, a line of evidence should group the whole set of homogeneous endpoints measured in all available experiments Higher tier studies may include purely experimental studies as well as modelling simulations, which may be considered 'in-silico experiments'. Within this section, the generic term 'experiments' is used without any distinction between these typologies (EFSA, 2018d).

A single result for an **endpoint** measured in one experiment is operatively defined as '**piece of evidence**', consistently with the definition reported in EFSA Scientific Committee et al. (2017). A line of evidence may consist of a single piece of evidence, of multiple pieces of evidence from the same experiment, or of multiple endpoints from several experiments. Thus, if the building blocks of the WoE are lines of evidence, the building blocks of each line of evidence are pieces of evidence (see Figure 13).

As for any scientific assessment, the WoE assessment should address a specific **problem formulation**, which may be translated into one or more **assessment questions**. This step in the present document is already described in Chapter 4 and must include considerations about the uses under assessment, the physical-chemical characteristics of the active substance, and potential interactions with other components of the product which may alter the environmental behaviour of the substance. However, the reader should be mindful that the problem formulation when assessing higher tier information may differ from the one formulated at the very beginning of the assessment i.e. at the lower tiers. This is because some aspects/questions of the risk assessment may already be addressed in the lower tiers, while newer ones may – under specific circumstances – arise.

EFSA Scientific Committee et al. (2017) considers that any WoE comprises three basic steps (see Figure 13):

- Assembling the evidence
- Weighing the evidence
- Integrating the evidence

A further, separate step of uncertainty analysis (see Figure 13) is also needed to take account of any other uncertainties affecting the overall assessment.

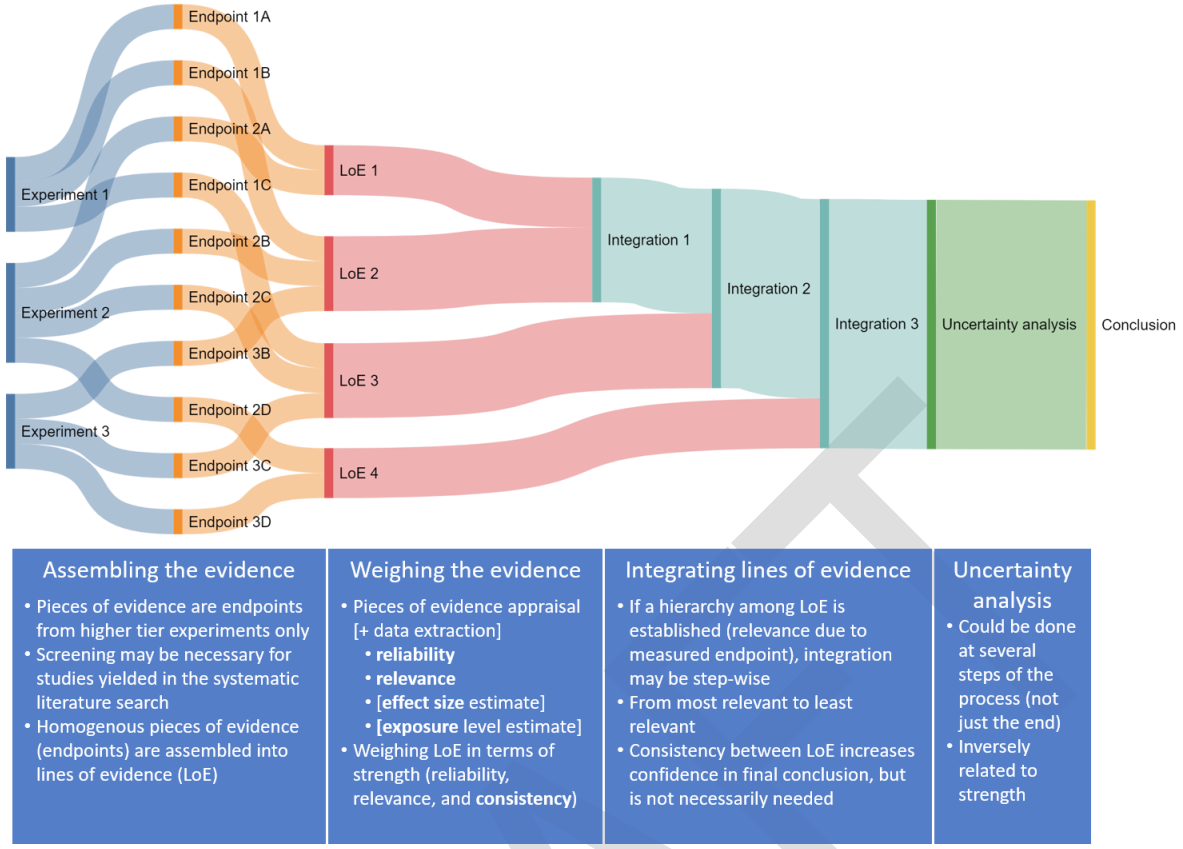


Figure 13: Summary of the WoE + uncertainty analysis process

10.6.2 Assembling the evidence

The first parts of this step, i.e. identifying and selecting evidence to include in the WoE assessment, is relatively straightforward in the context of the pesticide authorisation. The domain of the WoE is limited to the evidence submitted in the dossier, whether these are studies carried out by the Applicant or deemed relevant from the mandatory systematic review of the open literature. This also includes studies from the open literature which were flagged as relevant during the peer-review process. The definition of specific screening/eligibility criteria for the literature review should follow the recommendations of EFSA (2011).

Within the present guidance, it is recommended to include in the WoE uniquely evidence from higher tier studies (see Section 10.2). Evidence from lower and higher tier studies is generally not comparable, thus, any integration in common lines of evidence is discouraged. In principle, evidence from lower tier studies can be used for building additional lines of evidence. Nevertheless, tiered risk assessment schemes present by definition a clear hierarchy, where results from higher tiers generally override results from lower tier. Thus, it is expected that including evidence with low weight would not provide great benefit, when compared to the cost of integrating very heterogenous evidence. Other kind of information related to the specific uses under assessment, the physical-chemical characteristics of the active substance, and potential interactions with other components of the product, are better used to define the problem formulation (see Section 10.6.1) rather than being used as complementary lines of evidence. However, if for any reason this is not possible, it may be considered to include this information in the evaluation of the uncertainties (Section 10.6.5).

When assembling the evidence, one critical aspect is the grouping of pieces of evidence in lines of evidence. In some cases, this is relatively straightforward. However, care should be taken when similar but not exactly equal endpoints are measured and reported. For example, some experiments may report brood-related endpoints as 'area of the comb covered in brood', while other may report indexes related to brood development and termination. It is a responsibility of the risk assessor, depending on the type of WoE adopted (see Section 10.6.3) to judge whether pieces of evidence are sufficiently homogenous to be grouped in a single line of evidence.

10.6.3 Weighing the evidence

At this step of the WoE, risk assessors should evaluate the reliability, the relevance, and the consistency of pieces and/or lines of evidence.

Reliability, often referred to as 'internal validity', reflects the internal bias, i.e. potential errors in the conduct of an experiment that results in a conclusion which is different from the truth. The method for measuring an endpoint not being reliable/accurate is an example of source of internal bias. In some instances, reliability also includes considerations related to the precision in the measurements, which relates to the sample size/level of replication in comparison with the expected variability (e.g. power of the experimental design). However, in other cases, this is assessed separately (EFSA et al., 2020, EFSA PPR Panel et al., 2022b, EFSA PPR Panel et al., 2022a). Either way, this aspect should be considered.

Relevance, also referred to as external validity, reflect external bias, which affects the extent to which the study results are generalisable to the assessment question at hand. Relevance, in the present context, may be determined by two main aspects: 1) the ability of the measured endpoint to address the assessment question, which is further considered in the part describing the integration of the lines of evidence below; 2) the representativeness of the experimental settings, which may include considerations about environmental conditions, landscape structure, nature of the tested item (e.g. formulation type), treated crop(s), application technique and regime, etc. Relevance may or may not include considerations about the level of the exposure measured/estimated in the experiment in comparison with the exposure assessment goal (see more on this in Section 10.6.8).

When performing the appraisal of the individual pieces of evidence, it is good practice to extract the relevant information in a structured manner. This may include considerations about relevance and reliability in a form of critical appraisal tool, but it may also include information on: 1) the effect size observed in the experiment for the specific endpoint (see Section 10.6.7), 2) the exposure level and duration measured/predicted/imposed in the experiment (see Section 10.6.8)

Consistency reflects the level of coherence in the available evidence. While the assessment of reliability and relevance is applicable to single pieces of evidence as well as lines of evidence, consistency can only be appraised within (at this stage of the process) and between (at the next stage of the process) lines of evidence.

Pieces of evidence appraisal

In practical terms, this phase of the WoE starts with an evaluation of the relevance and the reliability of the available evidence. While it is acknowledged that some aspects of the evaluation will be applicable to all endpoints measured in the same experiment, other may substantially differ. Thus, it is strongly recommended that such an appraisal is performed at the level of the single endpoint measured in the experiment.

At this stage, it is worth distinguishing between the two main aspects identified above that contribute to relevance. In a line of evidence grouping homogenous endpoints, the ability of the measured endpoint to address the assessment question is a constant - and may in fact be determined *a priori* (see for example EFSA (2018c), EFSA (2018b), EFSA (2018a)). On the contrary, the representativeness of the experimental settings will likely differ between experiments, and thus should be evaluated carefully at this stage.

Weighing lines of evidence

Once the single pieces of evidence are appraised, the outcome can be used to appraise the **strength** of each line of evidence. Strength can be viewed as the degree to which the line of evidence allows achieving a conclusion for the assessment question. Overall reliability, relevance, and consistency, all contribute to the strength of the line of evidence (EFSA PPR Panel et al., 2022b, EFSA PPR Panel et al., 2022a).

A line of evidence composed of unreliable pieces of evidence cannot have a big strength, even if the relevance is high and the results are consistent across experiments – it will only provide weak evidence for the conclusion, possibly associated with big uncertainties (see Section 10.6.5). Similarly, a line of evidence composed by reliable and relevant pieces of evidence, which nonetheless provide an inconsistent picture in terms of results, will generally be considered inconclusive.

One of the issues which is often debated, is whether pieces of evidence classified as 'unreliable' should be included in the later stages of WoE. While there is little doubt that this evidence should have little to no weight in the assessment, it is here suggested that all pieces of evidence are included in the reporting also in the phases after the appraisal: this will increase the transparency of the process and would allow to estimate the impact that a different classification would have triggered as a part of the uncertainty analysis (see Section 10.6.5).

The relevance of the measured endpoint for addressing the assessment question plays a pivotal role in weighing the relative importance of the different lines of evidence. In fact, it may be appropriate in some case to establish an *a-priori* hierarchy among lines of evidence, depending on how much the measured endpoints are informative for the assessment question (EFSA, 2018c, EFSA, 2018b, EFSA, 2018a). Note that this relative importance depends once again by the assessment question defined in the problem formulation updated upon the lower tier risk assessment. For example, if a concern in the lower risk assessment is mainly identified for acute contact toxicity, the assessment question should mirror this. In such case a line of evidence summarising results for forager mortality is certainly more important that any brood assessment. This ranking is however completely reversed if the main concern identified in the lower tiers of the risk assessment is about brood development.

10.6.4 Integrating the evidence

If a hierarchy among lines of evidence is established, the integration may be performed in a step-wise manner, i.e. from the most relevant to the least relevant. If the most relevant line of evidence provides a conclusive answer to the assessment question with enough certainty, integration with other lines of evidence may not even be necessary. For example, if measurements of the attribute mentioned in the SPG (e.g. colony strength) provides a clear conclusive picture, no further assessment is necessary.

In many cases however, such a straightforward answer will not be available, and thus more lines of evidence will have to be considered. The way lines of evidence are integrated depends to a large extent on the chosen methods for performing the WoE. This section is purposefully not

prescriptive in this sense, and the assessor has the freedom to choose the method which is considered more appropriate for the specific case (see Section 10.6.3).

Nevertheless, as a general rule of thumb, having a consistent indication or concern or lack of it from several lines of evidence increases the confidence in the final conclusion. Thus, to some extent, the assessor should try to check the consistency across lines of evidence as well.

10.6.5 Uncertainty analysis

WoE assessment and uncertainty are closely related: the strength of any line of evidence and of the final conclusion are inversely proportional to the degree of uncertainty associated to it.

When performing an uncertainty analysis, the assessor should be mindful that, together with the uncertainty 'intrinsic' to the evidence, there may be additional uncertainty arising from the very same process used in weighing the evidence.

It is recommended that an uncertainty analysis is always performed together with any WoE assessment. This can be done in different ways, and uncertainty can in fact be considered at different steps of the process. Ideally, considerations about uncertainty should be included already when assessing reliability and relevance of the individual lines of evidence. However, it is even more important to clearly report the uncertainty associated to each line of evidence (See for example EFSA PPR Panel et al. (2022b), EFSA PPR Panel et al. (2022a)). Uncertainty can also be evaluated at the very end of the assessment process, just before reaching a conclusion (see for example EFSA (2018c), EFSA (2018b), EFSA (2018a)).

Uncertainty can be expressed in a purely qualitative manner, but attempts to quantify it to the maximum possible extent are encouraged. Reporting that make use of structured elements (e.g. tables) generally helps the reader in getting a better overview and is thus recommended.

10.6.6 Types of WoE

EFSA Scientific Committee et al. (2017) classifies methodologies used for WoE in four wide categories, based on where they stand in a scale from fully qualitative to fully quantitative, which to some extent also reflects a gradient in how formal the applied method is. These are: 'best professional judgment', 'causal criteria', 'rating', and 'quantification'.

The present guidance purposefully avoids being too prescriptive in the methods to be followed for the WoE, as different methods may be more appropriate in different cases. However, as a general rule of thumb, it is strongly encouraged that the whole process is formalised to the maximum possible extent since the beginning of the assessment, in order to increase transparency and reproducibility. At the very least, some formality is strongly recommended in the procedure used for appraising the individual pieces of evidence, in weighing the lines of evidence, and in presenting the available information in a structured way (e.g. tables and plots are preferred to long text descriptions). Establishing *a-priori* methods for integrating lines of evidence (other than ranking them by relevance of the measured endpoint) is generally less straightforward.

Different methods may also be applied to different steps of the same WoE. EFSA PPR Panel et al. (2022b), EFSA PPR Panel et al. (2022a) are examples of WoE integrating aspects of the category 'rating' (for the evidence appraisal), 'quantification' (for determining consistency within and between lines of evidence) and 'best professional judgment' (for integrating the lines of evidence, assessing uncertainty, and reaching a conclusion).

10.6.7 Determining effect sizes

Irrespectively of the method chosen for assessing and integrating the lines of evidence, whether this is more quantitative or qualitative (see Section 10.6.6), it is important that an effort is made to determine the magnitude of the effect observed for each measured endpoint in each available experiment i.e. each piece of evidence. In principle, both the magnitude and the temporal scale of the effect should be considered. However, for bees, the temporal scale of the effect is irrelevant in the SPG as currently defined (see Section 3.1).

'Effects' imply, by definition, a causality between exposure and an observed alteration of the measured endpoint, when compared to an untreated control. In reality, proving (or disproving) causality is always a complex task. As a starting point, it is recommended that deviations in both directions, i.e. favourable as well as adverse, are recorded and considered. This helps discriminating random variability from 'true' effects. Similarly, it is recommended that in a WoE approach, the assessor should not rely entirely on the outcome of statistical tests. Consistency in the results from several experiments with limited power may help overcoming inconclusive statistical evaluations performed on each single piece of evidence.

Most endpoints measured in bee higher tier testing present a remarkable temporal – and in some instances spatial – variability. This complicates the determination of the effect size. A possibility is to extract and report effect sizes in terms of ranges, rather than single estimates. However, the assessor should try to focus on the most relevant time in the experiment in consideration of the characteristics of the tested substance and on the expected exposure for the uses under investigation. A knockdown insecticide with short half-life applied to a flowering crop is expected to exert its effects soon after the application. On the contrary, effects on the number of adult solitary bees from the use of a persistent insect growth regulator may be seen only in the next generation, thus in some case (e.g. univoltine solitary bee species) one year after the treatment.

10.6.8 Exposure considerations

Exposure is a critical part of the risk assessment; EFSA Scientific Committee et al. (2017) suggests that the exposure estimates to be used in the risk assessment could be also included in a WoE as separate lines of evidence, later to be integrated with effect-related ones. In the context of the present bee risk assessment however, this practice is not supported. This is because a procedure is already in place for estimating - with a suitable level of confidence - the PEQ, i.e. level of exposure in line with the exposure assessment goal (ExAG). Such level should be used as a reference in assessing higher tier studies (see Section 10.5).

However, exposure in the individual experiments must be considered in the WoE. This may be done in two different ways. Exposure levels (magnitude as well as duration) can be considered in the relevance part of the appraisal of individual lines of evidence. If this approach is selected, the more the exposure level deviates from the PEQ and from the expected duration, the less relevant will the resulting piece of evidence be considered. The downside of this approach is that both high exposure and low exposure are considered less relevant, but without an explicit consideration of the influence played by the level of exposure on the measured endpoint. This issue could be addressed by using the second approach, i.e. considering exposure as a covariate within each line of evidence. If exposure levels influence the observed effect size in a sort of dose-response relationship, it may be possible to use this information to estimate the effect size at the PEQ. If a quantitative approach is selected, this can be performed following the principles of the meta-regression. Adopting such approach may also help in addressing the 'consistency' aspect within each line of evidence, e.g. by explaining difference in the measured effect sizes between experiments.

10.7 Ecological models for the support of higher Tier risk assessment

10.7.1 General suitability of ecological models for higher tier regulatory risk assessment

Ecological models for the higher tier refinement of pollinator risk assessment have, as any other modelling tool used in regulatory risk assessment, to prove their suitability, more precisely, they need to show appropriateness for supporting the regulatory question and their respective performance. One critical issue here is whether an ecological model provides an appropriate model complexity, meaning that the model needs to be sufficiently complex to match the relevant real-world processes and traits of pollinators in a landscape context. For a honey bee colony model, this could mean that internal processes in a beehive such as egg laying, brood care, recruitment, pollen and nectar inflow and consumption, are considered as simulated processes, with external influences, such as food quality and quantity, the performance of individual bees, the impact of varroa mites and infectious diseases, and other possible stressors, each accounted for in an appropriate level of detail. Additional processes linked to the interaction of the colony with the surrounding environment (e.g. foraging in the landscape) are generally extremely important aspects that determines an appropriate level of complexity. That does not mean that all processes need to be simulated based on first principles, since simulation models are, and need to be, always a simplification of reality. The key issue here is the potential simulation model's ability to cover these influencing factors in an appropriate way.

Furthermore, if the model is used to simulate exposure to a PPP and resulting effects, processes underlying these aspects should also be demonstrated to be considered appropriately. This includes for example the fate of the substance in the environment and within the hive, mechanisms that cause the exposure of the bees, the quantification of such exposure, and the link between the simulated exposure and the simulated effects. In this case, it is appropriate to refer to these models as ecological effect models. In fact, this appropriateness can be and need to be demonstrated by the comparison between model results and observed data, for example, from a relevant semi-field or field experiment. Based on a thorough definition of the corresponding environmental scenario, colony model predictions should show a good overlap with the dynamics of real observations, ideally with no or little calibration of model parameters, and should demonstrate matching with the most important patterns. For complex models, such as honey bee colony models, the demonstration of appropriateness can in general be performed on two levels. First, on the level of the general model, which means that parts and modules of the model can be shown to match with general experimental observations, e.g. the age structure of a honey bee colony matches with common observations. Second, the colony model needs to demonstrate that it can predict observations from a specific field trial within reasonable limits of prediction quality, and under consideration of the usually high uncertainty and variability of observations in the field.

Closely related to this second level of a check of appropriateness is the need to define respective environmental scenarios. For modelling bee population dynamics, the landscape context, i.e. type, amount and location of nectar and pollen resources, and weather conditions such as temperature, rain, irradiation are crucial input for the model, and need to be defined in the form of environmental scenarios. When using ecological models for extrapolation, additional environmental scenarios can be defined, e.g. to allow simulations of bee populations under less favourable conditions, e.g. in a low-resource landscape, or for other landscape composition or under different weather conditions. It is required that ecological models demonstrate reasonable simulations results for such a variety of environmental scenarios which can represent the different landscapes across Europe.

Depending on the intended use, in addition to the above-described need for an evaluation of the appropriateness of the model for describing a specific (set of) field experiment(s) could require additional work on scenario definition. If the aim of the application of the ecological model was for example to extrapolate to other areas of use (see below), carefully underpinned scenarios need to be selected that represent either well the entire variety of environmental conditions for the intended area of use of the compound (from which a desired percentile of probability of occurrence could be selected), or 'realistic worst case' environmental conditions in order to protect the bee population not under average conditions, but for e.g., 90 percent of all situations.

Last but not least, an appropriate, comprehensive and transparent documentation of an ecological model is an important criterion for its possible uptake for regulatory risk assessment. The use of an established protocol/format standard for the documentation of individual- or equation-based models such as ODD (Overview, Design concepts and Details) or TRACE (TRAnsparent and Comprehensive model Evaluation) is here considered as an important requirement. The documentation should allow and support the critical evaluation of the ecological modelling approach and the corresponding scenarios. A general framework how to evaluate ecological and ecotoxicological models for regulatory risk assessment of PPP was outlined in the EFSA scientific opinion on Good Modelling Practice (EFSA PPR Panel 2014).

10.7.2 Examples for supportive use of ecological (effect) models

Currently, there is no honey bee, bumble bee or solitary bee model established for the immediate use in regulatory risk assessment, including simulations of the exposure and effects to pesticides after spraying in agricultural fields. Therefore, the examples given below for the supportive use of ecological models in regulatory risk assessment should be considered as illustrative only. As mentioned in Chapter 3, the WG does not recommend the use of ecological models as higher tier refinement method in isolation, i.e. but with the support of semi-field or/and field data. The reason is, that there are still too many unknowns regarding the possible interactions between a pesticide application and its possible effects in a field situation while the models can only predict mechanisms which are appropriately considered and implemented. It is therefore a requirement that any higher tier assessment should be based on experimental data for the compound under assessment under realistic (semi-)field conditions, with modelling acting as a method to aid the interpretation of the experimental data and to extrapolate those to other situations.

10.7.2.1 *Extrapolation to untested conditions*

One possible way to utilise ecological models is extrapolating exposure observations from field tests to other locations or field situations. In a first step, an ecological effect model could simulate exposure as observed in an exposure field study with the aim to corroborate the results. In a second step, the model could be used to extrapolate the exposure situation to other landscapes and under different conditions e.g. weather, application times, landscape composition or other factors with the aim to refine bee colony/population exposure. In such case, it would be crucial to consider that a change in the environmental scenario should not only influence exposure, but may also alter colony dynamics and performance due to different conditions, e.g. a different food availability in the landscape, different weather, etc.

A special case concerning exposure refinement is the possible use of a simple foraging model for the calculation of exposure in a landscape context, which is proposed as a Tier 2 method (ref. Section 5.5.7 Landscape factor (LF)). While in general the same suitability criteria apply for such model application, the model's purpose would be to simulate nectar (or pollen) foraging, not colony/population dynamics, so that it would not appear immediately necessary to include complex colony/population-level or ecotoxicological simulations in such foraging models.

Nevertheless, such Tier-2 model would need to demonstrate its performance concerning nectar and pollen foraging against observations from field trials.

10.7.2.2 *Effectiveness of risk mitigation measures*

In a more general sense, higher tier ecological models for bees could be used to test the effectiveness of suggested Risk Mitigation Measures (RMMs) for different ecological and agricultural practice scenarios, and to help to prioritize them. For example, the effectiveness of flower strips, and the influence of their size (i.e. length/width) on the exposure of honey bee foragers in mass-flowering crops could be simulated. This has in a proof-of-concept study been shown already (Baveco et al., 2016), and more complex higher tier modelling studies could assess the effectiveness of RMMs in similar ways.

10.7.2.3 *Influence of additional stressors*

One of the most crucial questions when assessing the risks of pesticide applications to bee colonies/populations in field conditions, is whether bee colonies/populations are exposed and impacted by other stressors such as limited quantity and quality of nectar or pollen in monoculture landscapes, infectious diseases, climate change, etc. These other stressors might reduce the bees' resilience and lead to more severe impacts as compared to laboratory tests or in-field trials where some or several of these stressors can be excluded/controlled. In this case, ecological models can support higher-tier experimental studies, aiming at more holistic and realistic risk assessments including the consideration of multiple stressors at the landscape level.

10.7.2.4 *ApisRAM*

Recently, a honey bee colony model called *Apis*RAM was formalised (see Duan et al. (2022)) with the aim to determine the effects of multi-stressors (e.g. malnutrition, unfavourable body temperature, infectious agents and pesticides) on bee health and colony dynamics. The model is still in a developmental phase and will require additional calibration and testing before it can be used for the regulatory risk assessment of PPP. The model is an agent-based colony model which simulates colony-level behaviour emerging from actions and decisions made by individual interacting bees. The bees interact as well with, and are influenced by, their immediate environment (i.e. the in-hive physical and chemical state, comprising food stores as well as the environment outside the hive). The environment outside the hive is simulated within ALMaSS (the Animal Landscape and Man Simulation System) which is a spatially and temporally dynamic model comprising land use, farm practices, weather, crop growth, semi natural habitats and flower resources.

11 Metabolite

Regulation (EC) 1107/2009 requires that potentially harmful effects of an active substance and its metabolites on the environment shall be examined. According to the Regulation (EC) 283/2013 information shall be generated to permit an assessment of the impact of the active substance and its metabolites on non-target species, where they are of toxicological or environmental significance. Bees (honey bees, bumble bees and solitary bees) can be exposed to metabolites that are formed in pollen and nectar; the impact can result from single, prolonged or repeated exposure.

11.1 Method

The risk assessment scheme proposed in this Chapter focuses on exposure to residues following consumption of pollen and nectar that are contaminated by PPPs (dietary exposure).

A risk assessment for metabolites is triggered when:

- residues of metabolites are found at or above 10% TRR (Total Radioactive Residue) **and** 0.01 mg eq/kg (OECD, 2007) in residue studies in pollen and nectar or metabolism studies in primary and rotation crops OR

- residues of metabolites are found at or above 10% TRR (Total Radioactive Residue) **or** 0.01 mg eq/kg in residue studies in pollen and nectar or metabolism studies in primary and rotation crops, **and** their parent substance is of acute toxicity to bees (i.e. LD_{50} < 0.01 µg/bee).

Therefore, the first step is to assess whether a risk assessment is triggered for any identified metabolite in the available studies (e.g. plant metabolism studies). An overview of residue data, including recommendations on how the data should be assessed in order to get the relevant information on metabolites, is reported in Appendix D – to the Guidance Document. Guidance on how to conduct the risk assessment is given in Section 11.2.

11.2 Risk assessment for metabolites

When a metabolite requires further assessment based on residue data (i.e. ≥ 10% TRR and/or 0.01 mg eq/kg depending on the acute toxicity of the parent to bees), relevant information on the hazard and exposure of the metabolite to bees has to be provided.

11.2.1 Hazard characterisation

The data requirements for metabolites are the same as for the active substance (according to the Regulation (EC) 283/2013). Therefore, the following data are required for all metabolites formed in pollen and nectar (or plant matrices) at ≥ 10% TRR and 0.01 mg eq/kg OR at ≥ 10% TRR or 0.01 mg eq/kg if their parent is known to be acutely toxic to bees (i.e. LD_{50} < 0.01 µg/bee).

- Acute oral toxicity to bees
- Chronic toxicity to bees
- Honey bee brood study to determine effects on honey bee development and brood activity

The acute contact risk case to bees is not considered relevant as exposure to metabolites in nectar and pollen via contact is negligible.

At the lower tier effect level, the hazard is defined by the dose-response curve described by the D_{50j} and the slopej investigated in standard laboratory test, as for the parent (see Chapter 6).

For the screening effect tier hazard is not defined by laboratory tests with the metabolite but by toxicity data with the active substance and non-testing methods. This means that for the screening effect tier the toxicity data (LD_{50j}, slope) of the active substance, considering a 10-times higher or equal toxicity compared to the parent is used. For further details on the screening effect tier see scenario C below.

Choice of the (surrogate) endpoint (acute and chronic)

As mentioned above, when the metabolites require an assessment, applicant should provide the relevant information according to the data requirements (i.e. oral acute, chronic and larvae tests). Pending on the data available with the dossier, the risk assessor could consider the following scenarios:

Scenario A - Dossier complete: When acute and chronic toxicity data on bees are provided for the identified metabolite, then the relevant endpoints (i.e. LD_{50j}, slope) should be used for the risk assessment. If the metabolite is known to be less toxic than the parent (by at least a factor of 3), no further assessment is required as the risk is covered by the parent. If the metabolite is

known to be of the same or higher toxicity than the parent then a mixture toxicity risk assessment should be conducted. For further information on the mixture toxicity see Section 11 of the Guidance Document.

Scenario B - Dossier partially complete: When only acute toxicity data on bees are available for the identified metabolite, then the relevant acute endpoints (i.e. LD_{50j}, slope) should be used for the risk assessment (see scenario A). The chronic risk to bees can be estimated based on the available acute toxicity data. If it is shown that the metabolite acute endpoint is at least 10-times higher than the parent acute endpoint, then the metabolite can be assumed to be of the same chronic toxicity as the parent for both adult bees and larvae. Otherwise, further data have to be considered (see point 3 below) to estimate a chronic surrogate endpoint.

Scenario C - Data in the dossier are missing (screening effect tier): When no toxicity data on bees are available for the identified metabolite, or the available acute data do not allow a conclusive estimation of chronic toxicity, then the toxicity of the metabolite can be estimated based on other information. One approach to estimate the toxicity of a metabolite is to consider the results of toxicity studies conducted with the metabolite for other invertebrate species (e.g. aquatic invertebrates like *Daphnia* sp. or soil invertebrates like *Folsomia* sp. and *Hypoaspis* sp.). Acute studies are used to estimate acute bee toxicity, chronic studies are used to estimate chronic bee toxicity. If the data indicate that the metabolite is of the same or higher toxicity than the parent, a 10-times higher toxicity to bees compared to the parent should be used for the risk assessment and the endpoints should be considered as screening effect-tier. If the metabolite is of lower toxicity compared to the parent (by at least a factor of 3), the risk from the metabolite is considered to be covered by the parent.

Another approach to estimate the toxicity of a metabolite can be to take into account non-testing methods like e.g. QSAR or presence of the toxophore (see Appendix D – to the Guidance Document). When it is clearly demonstrated based on non-testing methods that the metabolite is not expected to be more toxic than the parent, the same toxicity as for the parent should be assumed for the metabolite risk assessment.
Overall, if the available data (e.g. QSAR or toxicity studies with other invertebrate species) indicate a higher toxicity of the metabolite compared to the parent, the submission of toxicity studies on bees with the metabolite should be considered by the applicant.

11.2.2 Exposure characterisation

For estimating the exposure to bees to metabolites the acute and chronic dietary exposure is considered for the adult honey bees and chronic exposure for the larvae.

The identification of the relevant metabolites is based on residue studies in pollen and nectar and plant metabolism studies; hence, measured residues of relevant metabolites are mostly not available. Therefore, the exposure characterisation should be followed for the screening and lower tier risk assessment, as described below. When appropriate measured residues of the relevant metabolites are available then the exposure characterisation should be based on the measured residues.

For the characterisation of the dietary exposure to the metabolite the exposure characterisation for the active substance is considered. Detailed information on the dietary exposure see Section 5 of the Guidance Document.

The exposure estimate for the active substance can be adapted taking into account the fraction of the parent to the metabolite (F_{met}). The F_{met} can be defined as the mass-to-mass fraction of the parent to the metabolite. The following equations should be used to determine the F_{met}:

$$F_{met} = F_{trr} \times M_{met} \div M_{par} \quad \text{[eq.1]}$$

Where:
F_{met}: fraction of the parent to metabolite (-)
F_{trr}: fraction of metabolite formed in the respective matrices (-)
M_{met}: molar mass of the metabolite - g
M_{par}: molar mass of the parent molecule - g

The fraction of the parent to the metabolite (F_{met}) is considered to determine the predicted exposure ($PEQ_{j,met}$).

$$Rint_{met} = PEQ_{j,met} = n \times AR \times Ef_{di} \times F_{met} \times \left(SV_{po,met} + SV_{ne,met}\right) \quad \text{[eq.2]}$$

Where:

$Rint_{met}$: Residue intake – ng/bee or ng/bee/day or ng/larva/developmental period in d
$PEQ_{j,met}$: Predicted Exposure Quantity due to dietary exposure, where the suffix *j* indicates one of the three risk cases, i.e. acute dietary, chronic dietary and larvae – ng/bee or ng/bee/day or ng/larva/developmental period in d
F_{met}: fraction of parent to metabolite (-)
Ef_{di}: exposure factor for dietary exposure (-)
AR: application rate of the parent molecule - g/ha
n: total number of applications before and during flowering (n_{be} + n_{du})
$SV_{po,met}$: shortcut value for pollen (µg/bee or µg/bee/day or ng/larva/developmental period in day
$SV_{ne,met}$: shortcut value for nectar (µg/bee or µg/bee/day or ng/larva/developmental period in day)

The number of applications should be considered for the calculation of the relevant exposure ($PEQ_{j,met}$). When the residue study was conducted according to the GAP, which means that the number of applications was already considered in the residue trial, then the number of applications has not to be considered in the equation (n = 1).

The SV_{met} for pollen and nectar is determined based on the standard parameters for the parent compound, i.e. landscape dilution, residues and consumption. However, it differs to the SV values determined for the active substance as parameters like e.g. DT_{50} in pollen and nectar and interval between the applications are not considered. Further, no difference between before and during flowering is made, the total number of applications is considered to be during flowering.
For the calculation of the SV values for pollen and nectar equations 3 and 4 should be considered.

$$SV_{po,met} = \frac{LF_{po} \times RUD_{po} \times CMP_{po}}{1000} \quad \text{[eq.3]}$$

$$SV_{ne,met} = \frac{LF_{ne} \times RUD_{ne} \times \frac{CMP_{SU}}{SN}}{1000} \quad \text{[eq.4]}$$

Where:

$SV_{po,met}$: shortcut value for pollen (ng/bee or ng/bee/day or ng/larva/developmental period in day;)
$SV_{ne,met}$: shortcut value for nectar (µgg/bee or ng/bee/day or µgg/larva/developmental period in day)

LF_{po}: landscape factor for pollen (-)
LF_{ne}: landscape factor for nectar (-)
RUD_{po}: residue unit dose of pollen – mg a.i./kg pollen
RUD_{ne}: residue unit dose of nectar – mg a.i./kg nectar
CMP_{po}: pollen consumption (mg pollen/bee or mg pollen/bee/day or mg pollen/larva/developmental period in day)
CMP_{su}: sugar consumption (mg sugar/bee or mg sugar/bee/day or mg sugar/larva/developmental period in day)
SN: sugar content of the nectar (kg sugar/kg nectar, i.e. -)

For the screening assessment worst-case assumptions were made for all parameters of the exposure estimation (e.g. Ef_{dj} = 1, worst-case sugar category), except for the fraction of the parent to the metabolite (F_{met}), the application rate (AR) and the number of applications (n).

For the risk assessment based on Tier 1 exposure estimates, the default values for each parameter of the exposure estimation should be used.

For further details on the calculation and the parameters used for the exposure estimation at the screening and Tier 1 see Chapter 5.

11.2.3 Risk assessment

Based on the available information on the hazard characterisation and the exposure estimate a risk assessment for metabolites should be conducted. The steps to be followed to properly assess the risk from exposure to each metabolite are outlined below.

For the combined evaluation of the exposure and hazard the SPGs for honey bees, bumble bees and solitary bees have to be defined.

For honey bees risk managers agreed on a magnitude dimension for the entire EU corresponding to a value of 10% as the maximum permitted level of colony size reduction following pesticide exposure. For bumble bees and solitary bees, a threshold of acceptable effect was not defined by the risk managers. Hence, there are no trigger values which would allow to interpret any quantitative lower tier outcome. However, based on this guidance, exposure estimation and hazard definition for bumble bees and solitary bees is possible and thus in further consequence a prediction of effects at colony/population level could be considered in a qualitative manner to tailor higher tier studies. For further details on the lower tier risk assessment for bumble bees and solitary bees see Section 7.3 of the Guidance Document.

- **Step 1**

Is there any metabolite formed in pollen and nectar (plant matrix as surrogate) at or above 10% TRR and 0.01 mg eq/kg OR at or above 10% TRR or 0.01 mg eq/kg if their parent is of acute toxicity to bees (i.e. LD_{50} < 0.01 µg/bee)?

No: No relevant metabolites are formed. No further consideration is needed.

Yes: Metabolite risk assessment is triggered. Go to Step 2

- **Step 2**

Conduct a screening risk assessment for bees (honey bees, bumble bees and solitary bees) based on exposure estimated at the screening level and the selected toxicity endpoints (see section above).

As for the active substances, also for the metabolites it is recommended to follow the approach based on combined toxicity for risk assessment. Detailed information on the relevant exposure scenarios to be considered in the screening assessment see Section 7 of the Guidance Document.

Outcome of the risk assessment:

PE_{SPG} ≤ 10% Low risk is concluded

PE_{SPG} > 10% High risk due to exposure to the metabolite cannot be excluded. Go to Step 3.

- **Step 3**

Conduct a risk assessment based on Tier 1 exposure estimates for bees (honey bees, bumble bees and solitary bees) following the same approach outlined for the screening assessment (see Step 2)

Outcome of the risk assessment:

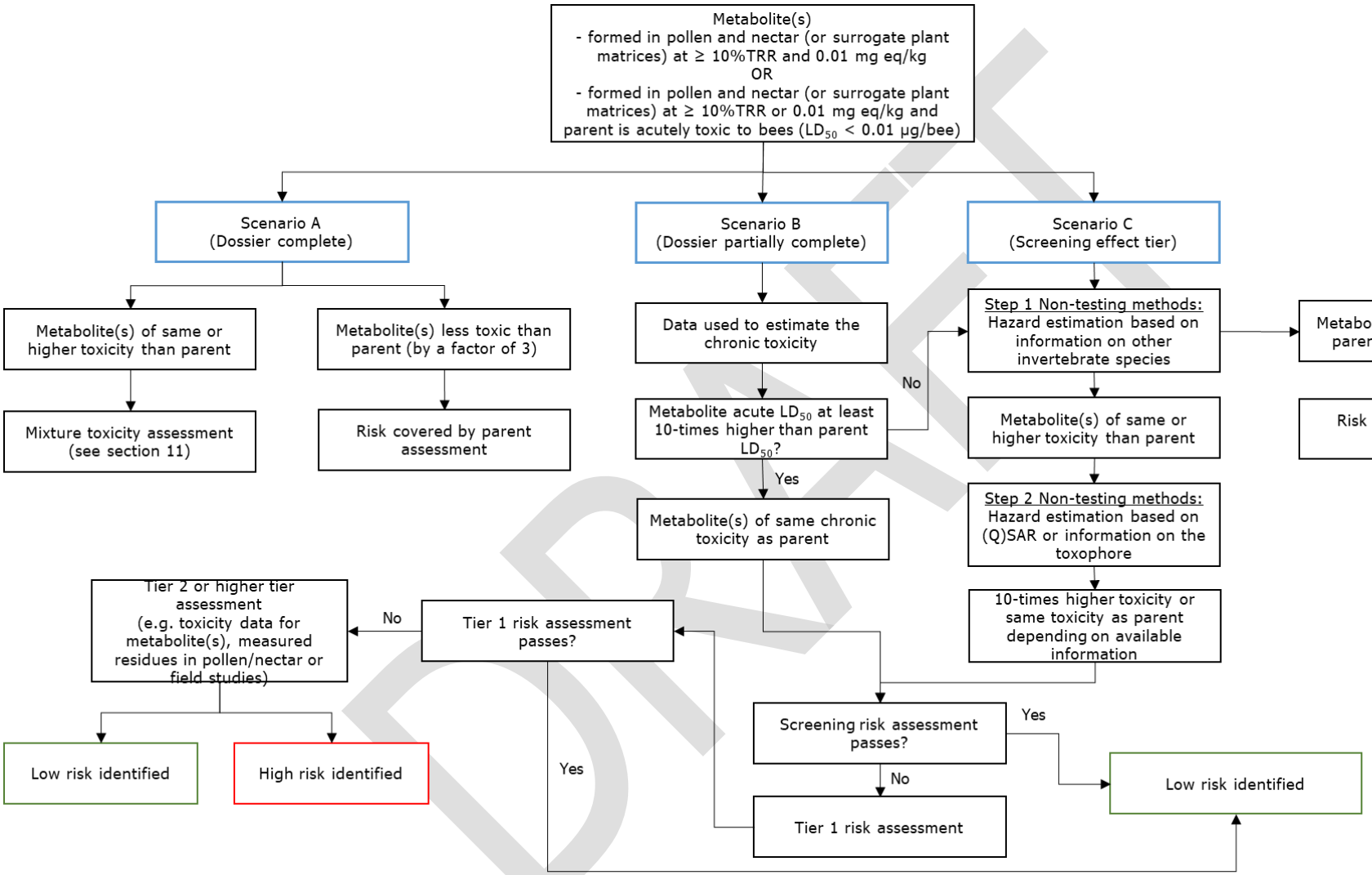
PE_{SPG} ≤ 10% Low risk is concluded

PE_{SPG} > 10% High risk due to exposure to the metabolite cannot be excluded. Go to Step 4.

- **Step 4**

At a first place it should be considered taking into account the refinement of the exposure at Tier 2 (i.e. measured residues or residue decline in pollen and nectar, for further details see Annex A and B of the Guidance Document) or of the hazard (toxicity studies with the metabolites to be requested, if only screening effect tier was available). It might also be useful to conduct residue trials in pollen and nectar showing that the metabolite is not relevant considering the critical GAP (see Section 11.3). Detailed information on the tier 2 risk assessment see Chapter 7.

In a next step higher tier studies (i.e. field effect studies) should be conducted to address the risk to bees. Field effect studies conducted with the active substance, or a relevant formulation, can be considered to address the risk from metabolites. In general, higher tier effect studies conducted according to the actual representative use of the active substance should also cover the risk from metabolites; hence, confirmation of presence of the metabolites is not needed. For detailed information on higher tier studies see Section 8 of the Guidance Document.



4090

4091 **Figure 14:** Flowchart for metabolite risk assessment

12 Mixtures

In this Chapter, an approach is proposed to address the risk of mixtures for honey bees, bumble bees and solitary bees with focus on technical mixtures (products) containing more than one active substance and their co-formulants undergoing an authorization procedure, solely. However, the concept and scheme proposed here is generally also applicable to other mixtures e.g., spray tank mixtures where different types of PPPs are mixed that may involve certain combinations of concern with potential for synergistic effects (e.g. pyrethroids insecticides and azole fungicides). Such combinations of substances are not expected within a technical mixture of active substance(s) and their co-formulants undergoing an authorization procedure with the Regulation (EC) 1107/2009, those mixtures being registered for a specific use/target. Tank mixtures are therefore not within the scope of this guidance but still they may result in enhanced toxicity (or even when PPPs are applied in close temporal proximity). When information is available for such combinations of substances, the concept and scheme proposed within this Chapter may also apply.

12.1 Legal requirements

The Regulation (EC) No 1107/2009 requires in Article 29 that "interaction between the active substance, safeners, synergists and co-formulants shall be taken into account" in the evaluation and authorization of a PPP. This explicitly refers to marketed PPP, which are, by origin, technical mixtures containing one to several active substances, plus, typically, several co-formulants. Furthermore, the standard data requirements for PPP (Commission Regulation (EU) No 284/2013) do request "any information on potentially unacceptable effects of the plant protection product on the environment, on plants and plant products shall be included as well as known and expected cumulative and synergistic effects".

The approach builds on existing methods and scientific experience in assessing chemical mixtures. In most cases, mixture effects are based on adding up the doses for common effects to estimate the overall risk. However, sometimes, the chemicals 'interact', meaning their toxicity increases or decreases. Interactions need checking particularly if toxicity increases.

12.2 Risk assessment for mixtures

12.2.1 Defining the hazard

For the RA of mixture under Regulation (EC) No 1107/2009, applicants and risk assessors should define which approach is more suitable for the hazard assessment. Two options are considered possible that involve measured ("whole mixture" approach) and calculated mixture toxicity ("component based" approach) pending on the time frame of the exposure and the residue dynamic of the various components of the mixture. The use of calculated mixture toxicity should be considered whenever justified (i.e., *a priori*, no synergistic effects) and when this is the only possible approach (e.g., mixture composition of a.s. is different in the formulation than expected in the environment or experimental testing is technically not feasible). On the basis of the mixture toxicity (measured or estimated) selected for each risk case (i.e. acute contact, acute dietary, chronic-dietary and larvae-dietary), a combined risk assessment can be conducted for each bee group, in line with the approach explained in Chapter 7.

In case of measured toxicity, the selection of the relevant hazard parameters ($LD_{50,j}$, $slope_j$) will follow the same rules as explained in Chapter 6.

There may be situations where the dose response of a mixture is not symmetric, e.g., in case the different compounds have very different slopes (refer to example below). Whenever this is the case, the dose-response will be characterized by a shallower slope at lower doses and by a steeper slope at higher doses. The opposite situation is not expected to occur.

As mentioned in Chapter 6, once an LD_{50} is fixed, a shallower slope would predict a higher mortality (and thus represent a worst-case) than a steeper one for doses below the LD_{50}. The opposite is true for doses above the LD_{50}. Particularly for the honey bee risk assessment, the focus should primarily be on doses causing up to 10% mortality[10], as a PEQ_j causing a higher mortality would immediately result in a high risk.

In view of the two points above, for the measured mixture toxicity, it is important that the selection of the hazard parameters to be used in the risk assessment would ensure that particularly the left-hand part of the dose-response curve is correctly described, or in other words, that mortality is not underestimated at the lower doses.

For calculated mixture toxicity, the potential asymmetry of the estimated curve renders the description of the dose-response less straightforward.

Indeed, when using slope data in dose addition (DA) calculation, the 'predicted' dose-response for the mixture is not necessarily a conventional symmetric curve, even if the dose-response curves of the individual substances are. In such case, the dose response curve of a mixture may not be described as a standard 'log-logistic' model (i.e. as it is done for single active substance), which means that it cannot be described by a simple combination of LD50 and slope.

Under a DA approach, for a mixture of *n* components, a specific $LD_{x,mix}$ resulting in an effect level *x* is calculated as follows:

Equation 1: $$LDx_{mix-DA} = \left(\sum_{i=1}^{n} \frac{p_i}{LDx_i}\right)^{-1}$$

Where:

n: number of mixture components

i: index from 1…n mixture components

p_i: the i^{th} component as a relative fraction of the mixture composition (note: $\Sigma\ p_i$ must be 1)

LDx_i: dose of component i provoking x% effect

[10] For bumble bee and solitary bee, such threshold is not available, but it is considered highly unlikely that, once a quantified SPG is available, the magnitude of acceptable effect will be above 50%. Thus, a shallower dose-response would still represent a worst-case.

This means that, when the dose-response relationships and the relative proportions of the n components of the mixtures are known, it is possible to calculate the $LD_{x,mix}$ for a range of effect levels. This allows describing the dose-response of the mixture in a rather precise way.

The figure below provides an illustration for a 1:1 mixture of substance A and substance B, with the same LD_{50} of 10 (generic unit). The (red) curve describing the mixture is closer to Substance A in the left part of the plot, and then closer to Substance B in the right part. This is because the difference in the slope makes substance A the driver of the effects at lower doses and substance B the driver at higher doses. The asymmetry of the resulting curve makes it unsuitable to be described by a common log-logistic model, and therefore cannot be described by a simple combination of LD50 and slope.

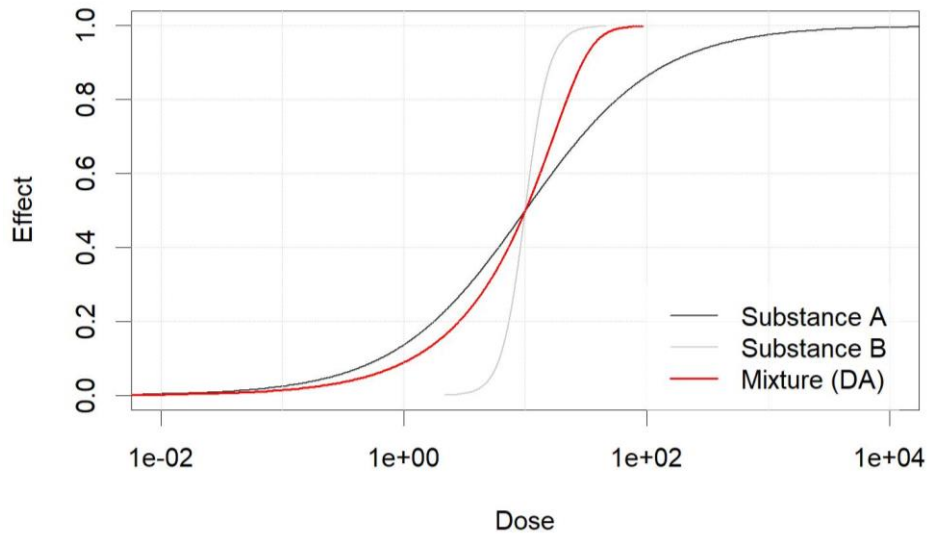


Figure 15: Illustration of an asymmetric dose response curve resulting from a Dose Addition model applied to a mixture of two substances with equal LD_{50} and different slopes.

This asymmetry is expected mainly when the mixture is composed by substances whose slope is very different, while the ratios of their LD_{50}s to their relative proportion in the mixture are in a similar range (how similar exactly would depend on the slopes). If these two conditions are not verified i.e. when the curves are more or less parallel and/or when the ratios of their LD_{50}s to their relative proportion in the mixture are very different (i.e. in presence of a driver substance), the mixture dose-response remains similar to a log-logistic model output.

Even when the dose-response cannot be approximated by a log-logistic model, an estimation of the effect of the mixture due to a specific level of exposure can still be made. To do this, it is proposed to calculate the mixture dose needed to achieve a suitable number of effect levels (i.e. from 1% to 99%) by using equation 1. Some minimum interpolations between the points would be needed. The resulting curve cannot be explained as a standard 'log-logistic' model, which means that it cannot be described by a simple combination of LD_{50} and slope, but can still be used to predict with reasonable accuracy, the effects caused by a certain exposure level.

12.2.2 Defining the exposure of the mixture to be assessed

The basic concept of the risk assessment for bees is that they are exposed to residues of active substances in the environment, e.g. via their food. Thus, the following steps refer to

4196 the assessment of effects from exposure to a mixture of active substances (and possibly also
4197 toxic co-formulants) in the environment resulting from use of a formulation.

4198 An LD50 for a mixture of active substances calculated assuming dose additivity can be
4199 conceived as an LD50 of a single virtual compound. It is thus deemed the most logical
4200 approach to also base the exposure side of the risk assessment on the same assumption.
4201 Content in the formulation and application rate per hectare should thus be expressed in terms
4202 of this virtual compound. The default residue dynamics used for both pollen and nectar (for
4203 each individual substance) are also applicable to the mixture as a single virtual compound.

4204 However, if the dietary exposure assessment is be refined using specific environmental fate
4205 data for individual active substances, the composition of the residues might be changed as
4206 compared to the original mixture. The exposure level $Rint_{mix}$ can be calculated as follows.

4207 Equation 2:

4208 $$Rint_{mix} = \sum_{i}^{n} Rint_i$$

4209 With:

4210 $Rint_i$ = Exposure expressed as "residue intake" of active substance *i*. This parameter
4211 corresponds to the exposure part of the dietary model that was designed for individual
4212 substances and may include substance specific residue dynamics (e.g. refined RUD, refined
4213 dissipation half-life, etc.).

4214 It should be carefully checked whether metabolites of ecotoxicological relevance (see Section
4215 11) have to be included into the $Rint_{mix}$ or not. For an initial screening approach, it may be
4216 assumed that the $Rint_i$ of all a.s. present in the formulation and the relevant metabolites that
4217 are formed from these substances will occur at the same moment and are not separated in
4218 time (i.e. worst-case $Rint_{mix}$). If on this basis the risk is not excluded, in a subsequent step,
4219 more detailed consideration of the predicted exposure patterns in time should be undertaken
4220 for mixture RA. The "metabolite intake" and the hazard parameters (whether surrogate or
4221 measured) should be based on the risk assessment scheme for metabolites available in Section
4222 10. This applies for both acute and chronic exposures.

4223 If the risk assessment is based on experimental toxicity data for the formulated product, no
4224 differentiation according to environmental fate parameters of individual active substances is
4225 possible. In principle, the concept of the single virtual compound could also be applied to
4226 calculate time-weighted average concentrations (TWA) for mixtures or formulations. The
4227 default fate parameters used for individual substances also apply here.

4228 Only active substances were used in the dataset from which the default fate parameters of
4229 the dietary model are derived, then uncertainty remains on the real fate of other co-formulants
4230 present in the formulation that is applied. Co-formulants may in some cases dissipate slower
4231 than the active substances and would not be covered at the screening and tier 1 risk
4232 assessment level. In absence of specific data, uncertainty remains on the actual exposures of
4233 bees to these compounds.

4234 It is noted that the above considerations only concern dietary exposure, and not contact
4235 exposure, the latter being relevant only on an acute timeframe and estimated on the basis of
4236 a single application. Indeed, for acute contact (i.e. overspray), bees are always exposed to
4237 the formulation 'as is', without any shift in its composition.

12.2.3 Risk assessment scheme

A detailed step-wise decision scheme is proposed below.

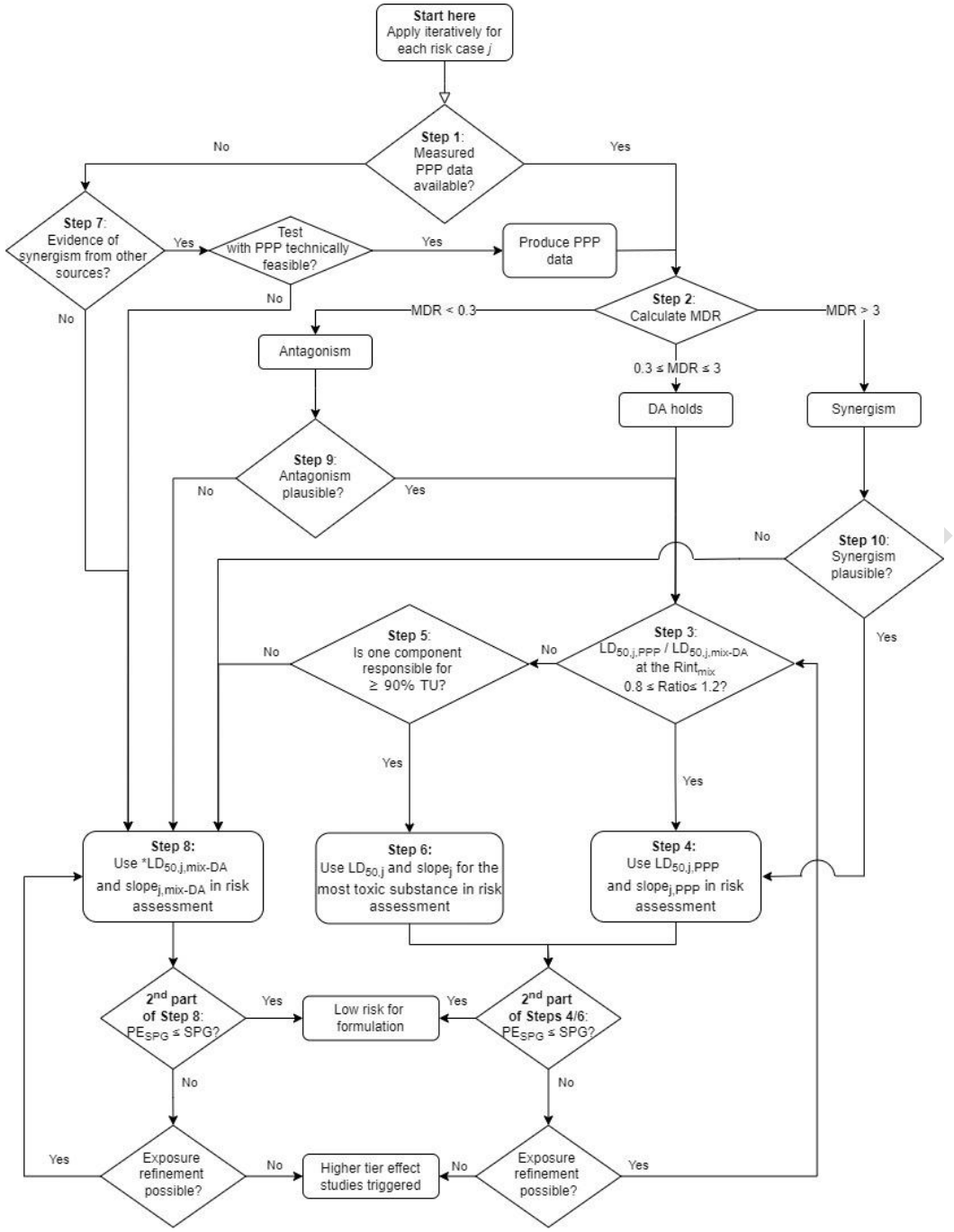


Figure 16: Workflow illustrating the risk assessment scheme for mixtures. *$LD_{50,j,mix\text{-}DA}$ may need to be corrected by an appropriate MDR from other risk cases/species when synergism is plausible (see text in steps 7-8).

The selection of hazard parameters (LD50j, slopej) for each risk case (i.e. acute, chronic, larval development) is based on the following scheme (to be applied reiteratively for each risk case, see Figure 16). Each of them can then be used in the combined risk assessment, either if based on measured or estimated toxicity.

Step 1. Are measured toxicity data (LD50j) available for the given risk case (typically, when PPP contains only one active substance, chronic data may be available only for the active substance, see Section 6)?

No, (Only data for the a.s. ($LD50j_{a.s.}$) are available): Go to 7
Yes, (both data for formulation ($LD50j_{PPP}$) and active substance ($LD50j_{a.s.}$) are available: Go to 2.

Note that data on slope are not required at this stage (endpoint selection only requires a comparison between measured and calculated LD_{50}s, it does not require an assessment of effects based on estimated exposure levels).
Note for acute (dietary and contact), data are generally available for both PPP and active substance(s).

Step 2. Check the plausibility of the calculated mixture toxicity LD50,j,mix-DA assuming dose addition (DA) against the measured formulation toxicity (LD50,j,PPP) on the basis of the mixture composition of the active substances in the formulation by means of the Model Deviation Ratio (see equation 3).

Notes:

In order to determine if the active substance may act more (i.e. synergistically) or less (i.e. antagonistically) than expected by DA, a comparison of the calculated $LD50j_{mix\text{-}DA}$ for the mixture composition of active substance in the formulation *versus* measured $LD50j_{PPP}$ endpoints is informative.

Thus, the first phase consists in estimating the mixture toxicity assuming the principle of dose addition ($LD_{50,j,mix\text{-}DA}$, see Equation 1).

Note that at this stage, the intent is only to compare the calculated and the measured LD50j of the PPP, and not to confront those values to an exposure estimate (Rint). Thus, the slope is not necessary at this stage.

This comparison may also indicate that other co-formulants not included in the calculation contribute to the overall mixture toxicity in an appreciable way. When this is the case, they can be included in a refined calculation (if the respective single-compound toxicity data are available). Possible outcome of the MDR calculation is the following:

$$MDR = \frac{\mathrm{LD50jmix} - \mathrm{DA}}{\mathrm{LD50jPPP}}$$

- **0.33≤MDR≤3.** The observed and calculated LD50j are considered in agreement if the MDR is between 0.33 and 3. This convention is in line with the recent EFSA recommendations related to pesticide risk assessment (Pesticide Peer Review Meeting 185, 9–12 October 2018[1]). In relation to 'when a formulation should be considered more toxic than the active substance', the proposal was to account for a difference of a factor of three, as recommended in the guidance from the Directorate-General for Health and Food Safety (SANCO/10597/2003 rev. 10.1) (European Commission, 2012) on the equivalence of batches and in the aquatic guidance (EFSA PPR Panel, 2013). Thus, if the calculated MDR lies between these two values, it is considered that the DA

hypothesis holds. This factor was agreed to be applied consistently to Tier 1 studies for all groups of non-target organisms including bees.

- **MDR is > 3.** More-than additive (i.e. synergistic) mixture toxicity is indicated if the MDR is > 3.
- **MDR is < 0.33.** Less-than additive (i.e. antagonistic) mixture toxicity is indicated if the MDR is below 0.33.

A careful interpretation of the MDR is mandatory, especially if not all components that potentially contribute to the observed mixture toxicity (e.g. co-formulants) have been considered in the DA calculation. Care should also be taken that the counter-checking of measured and calculated LD50j refers to the same basis, that is, the relative proportion of mixture components must be consistent (e.g. to the sum of active substances of a given PPP if co-formulants are not included in the DA calculation).

If MDR = 0.33–3 (DA approximately holds for the mixture): Go to 3
If MDR < 0.33 (mixture less toxic than DA): Go to 9
If MDR > 3 (mixture more toxic than DA->potential synergism): Go to 10

Step 3. Check whether the mixture composition in the formulation study giving the measured mixture toxicity ($LD50j_{PPP}$) in terms of the relative proportions of the individual active substance is similar to the mixture composition at the $Rint_{mix}$. As a direct comparison on the basis of the relative proportions of the a.s. at the $LD50j_{PPP}$ with the relative proportion at the $Rint_{mix}$ is not informative as such, the comparison is done based on calculated mixture toxicity (assuming DA) for both mixture compositions. Therefore, calculate $LD50j_{mix\text{-}DA}$ for the mixture composition of the a.s. at the $Rint_{mix}$ and compare with the estimate calculated for the formulation (as already done in step 2 above).

Note: this step is necessary only if a shift in relative proportions of the individual compounds is expected i.e. if substance specific fate characteristics (e.g. DT50) are used. Otherwise the default fate parameters apply to all compounds and no change is then expected in their relative proportions. This check between mixture compositions is considered necessary in case the tier 1 risk assessment fails and tier 2 is needed which may require the refinement of fate parameters for acute and chronic risk assessment (it may be anticipated that refined DT50 will have an impact on both the maximum expected residue level in case of multiple application (and thus on the acute risk) as well as the average residue level over a prolonged exposure (and thus on the chronic risk). The comparison is done based on calculated mixture toxicity (assuming DA) for both mixture compositions, that is, a calculation of $LD50_{mix\text{-}DA}$ for the mixture composition of the a.s. at the $Rint_{mix}$ and comparison with the respective estimate calculated for the formulation. The relative proportion of a.s. is considered sufficiently similar if the outcome of these calculations deviates less than 20% (i.e. as already recommended in the guidance document for aquatic organisms). Hence, if $LD50_{PPP}$ (proportion of a.s. as contained in PPP) divided by $LD50_{mix\text{-}DA}$ (proportion of a.s. at $Rint_{mix}$) yields a value between 0.8 and 1.2, a direct comparison of $Rint_{mix}$ with the $LD50j_{PPP,\ slopejPPP}$ is feasible. If the mixture composition differs more profoundly, the measured data cannot be used directly; however, they might be used to justify the use of the calculated approach to perform the mixture RA. Note that the results depend on the exposure scenario and the toxicity of each substance for each organism. Therefore, the check has to be performed for each application scheme in the GAP and for each organism separately.

If $LD50j_{mix\text{-}DA}$ (a.s. in PPP)/$LD50j_{mix\text{-}DA}$ (a.s. in $Rint_{mix}$) = 0.8–1.2 (mixture similar): Go to 4
If not (mixture not similar): Go to 5

Step 4. Use the measured mixture toxicity ($LD50j_{PPP}$ and $slopej_{PPP}$ for the risk case of concern) and proceed to the RA.
Note: in case synergism is plausible (following step 10) this can only be accounted for by assuming that the relative proportion of the mixture components is fixed. A change in the mixture composition would likely be reflected in another, unknown, MDR value. Thus, when synergism is plausible, exposure refinements can only be considered if these are not causing a shift in the relative proportion (e.g. if refined DT50s are available for different components of the mixture, the worst-case among them should be applied to all components).

PE_{SPG} ≤ SPG: Low risk
PE_{SPG} > SPG: low risk not demonstrated/check refinement options

Note that PE_{SPG} represents the combined effects of all the risk cases taken in consideration in the RA. To comply with the SPG, any risk case can be refined independently of the others and thus a refinement on an other risk case may suffice.

Step 5. Check whether one mixture component clearly drives the toxicity if considering the measured mixture toxicity ($LD50j_{PPP}$), that is, does the largest part of the sum of toxic units (TU) calculated for the formulation (≥ 90 %) comes from a single a.s. (TU_i)?

Note: if the toxicity of the mixture is largely explained by the toxicity of a single a.s., a sufficient protection level might be achieved by simply basing the RA on the toxicity data for that single "driver". Hence, where DA provides a reliable (or worst-case, for antagonism) estimate of the toxicity of the given mixture ($LD50j_{PPP}$) and the largest part of the sum of toxic units (i.e. ≥ 90 %) calculated for the measured mixture toxicity ($LD50j_{PPP}$) by Equation 4 comes from a single a.s., it can be concluded that this component drives the overall mixture toxicity.

Equation 4:

$$\sum_{i=1}^{n} TU_i = \sum_{i=1}^{n} \frac{c_i}{LDx_i}$$

Yes, (single "driver" of mixture toxicity identified): Go to 6

No, Go to 8

Step 6. Use the single-substance data ($LD50j_{a.s.,\ slope}$) for the identified "driver" (for the risk case of concern) and proceed to the RA.

PE_{SPG} ≤ SPG: Low risk

PE_{SPG} > SPG: Low risk not demonstrated/Check single-substance refinement options.

Note that PE_{SPG} represents the combined effects of all the risk cases taken in consideration in the RA. To comply with the SPG, any risk case can be refined independently of the others and thus a refinement on the driver might not be necessary (e.g. if a driver is identified for the chronic RA, a refinement on larvae may suffice to comply with SPG).

Step 7. Is there evidence that synergistic interactions between mixture components might occur (e.g. based on toxicological knowledge from literature or from counter-checking measured and calculated mixture toxicity for risk cases) which cannot be ruled out for the given endpoint with sufficient certainty?

Note: if synergistic effects cannot be excluded, the risk assessment should preferably be based on measurements, as synergistic interactions are not predictable by DA nor by other concepts such as independent action /response addition. If experimental testing of the mixture is no option (e.g. for technical reasons) for certain species and endpoints, but synergism is known form other studies, the RA may be performed by **adjusting the calculated $LD_{50,j,mix\text{-}DA}$ by the MDR** obtained from other risk cases/species.

Yes (mixture toxicity calculation not feasible): Measured mixture toxicity data required for RA

- If measured mixture toxicity becomes available: Go to 2
- If measuring the mixture toxicity is not technically feasible, but a reliable MDR is available from other risk cases/species, adjust the calculated $LD_{50j,mix\text{-}DA}$ by the MDR and go to 8

No (mixture toxicity calculation feasible): Go to 8

Step 8. Use the calculated mixture toxicity ($LD50j_{mix\text{-}DA,\ slope}$) for the risk case of concern) and proceed to the RA.

Note: in case synergism is plausible this can only be accounted for by applying an appropriate MDR (see step 7). The use of the MDR however works if the relative proportion of the mixture components is fixed. A change in the mixture composition would likely be reflected in another, unknown, MDR value. Thus, when synergism is plausible, exposure refinements can only be considered if these are not causing a shift in the relative proportion (e.g. if refined DT50s are available for different components of the mixture, the worst-case among them should be applied to all components).

PE_{SPG} ≤ SPG: Low risk

PE_{SPG} > SPG: Low risk not demonstrated, check single-substance refinement options

Note that PE_{SPG} represents the combined effects of all the risk cases taken in consideration in the RA. To comply with the SPG, any risk case can be refined independently of the others and thus a refinement on an other risk case may suffice.

Step 9. Carefully recheck the apparent antagonism as observed in the measured mixture toxicity data ($LD50j_{PPP}$) regarding potential impacts of the default assumption of DA. Does the apparent antagonism holds?

Note: If plausible toxicological explanation for this apparent antagonism can be provided (e.g. special feature of the formulation type), the RA should be based on the measured toxicity. Otherwise, the calculated mixture toxicity is a better option. No correction for MDR is needed, as the calculated mixture toxicity represents a worst-case.

Yes (measured mixture toxicity not plausible): Go to 8

No (measured mixture toxicity plausible): Go to 3

Step 10. Carefully recheck the apparent synergism as observed in the measured mixture toxicity data ($LD50j_{PPP}$) regarding potential impacts of heterogeneous input data (a.s.) and of co-formulants ignored in the DA calculation. Does the apparent synergism holds?

Note: If plausible toxicological explanation for this apparent synergism is available, the RA should be based on the measured toxicity. Otherwise the calculated mixture toxicity is a better option.

Yes: Go to 4
No: Go to 8

13 Risk mitigation measures

13.1 Introduction

If a high risk to honey bees, bumble bees and/or solitary bees is indicated at lower tier (i.e. Tier 1, Tier 2) or higher tier, an option to refine the risk is to consider risk mitigation measures to reduce the exposure of bees. These can be proposed by the applicant with the submission of the dossiers.

The aim of this Chapter is to summarise currently used risk mitigation measures as well as ongoing projects on this topic to improve and harmonise the risk mitigation measures applied. In addition, issues developing/applying such measures are discussed in detail.

In applying/developing risk mitigation measures, it is important to define key elements such as flowering. Presented in Table 1 are some suggestions regarding possible definitions of flowering, MSs may wish to develop their own definitions.

Table 28: Definitions that may be useful in determining when risk mitigation can be applied

	Definition
General flowering (bloom)	Flowers in which the stamen or pistils are visible. A crop is considered a flowering crop when the first flowers are open (BBCH 60)
Flowering crop - orchard	An orchard is considered a flowering crop when more than 1% of the flowers in an orchard are flowering.
Flowering crop - field crops	The crop is considered a flowering crop when more than two plants (crop and/or weed plants) per square meter are flowering.

Flowering crop - flower bulbs/bulb flowers	A crop is in flower when more than 1 % of the plants in a field is flowering. In Dutch agricultural practice, this means that a crop is considered to be flowering when more than two plants per linear meter of a field are flowering.
Flowering weeds	When there are five or more open weed blooms per m2 on average for the area being measured.

13.2 Risk mitigation measures

Risk mitigation measures can be classified into several broad categories with the aim to reduce the exposure to bees.

- Specific mitigation measures (quantifiable) are targeted actions which are needed to mitigate an identified risk due to pesticide exposure. It is important to note that specific mitigation should sufficiently reduce the risk to a level which is considered to be in line with the agreed SPGs (i.e. a low risk using a refined exposure estimate considering the mitigation must be demonstrated). As such any suggested mitigation must be accompanied by an appropriate risk assessment for which additional data may be needed. Specific mitigation measures can be proposed by the applicant within the risk assessment process.
- Generic mitigation measures (unquantifiable) are those actions which are undertaken to manage the risk to bees (e.g. risk phrase SPe8). Generic mitigation measures can be considered as risk management options within the decision-making process for the approval of substances at EU level and/or the authorisation of PPPs at MS level.

In general, by implementing risk mitigation measures repercussions on other sections, especially on efficacy, have to be taken into account. By considering measures like the modification of the GAP or the use of drift reducing techniques an acceptable level of efficacy has to be ensured.

13.2.1 Specific risk mitigation measures

As mentioned above specific mitigation measures should allow to reduce the exposure estimation and consequently to refine the PEQj values at the tier 1 or tier 2 (see Sections 3 and 5).

For example, applicant can propose drift reduction technique for spray applications or a higher quality of treated seeds to reduce the exposure estimation for the field margin /adjacent crop scenarios or specific technologies to reduce the exposure of flowering weeds in the field. When a high risk is identified e.g. for the treated crop scenario, no specific mitigation measures can be considered to reduce the exposure (i.e. to recalculate the PEQj), however applicants can also propose with the submission of the dossiers, GAPs that would allow to reduce and/or avoid the exposure such:

a. Restrictions on growth stage e.g. outside the flowering crop period
b. Restrictions to permanent greenhouse
c. Reduction of application rate/number of application/interval between applications

It should be noted that changes to the GAP during EU level assessments are not allowed and therefore applicants should carefully consider their range of selected GAPs when developing a dossier.

13.2.2 Generic risk mitigation measures

With regard to the protection of bees, specifically to managed honey bees, Regulation (EU) No. 547/2011 provides a set of 'risk phrases' aimed at reducing the exposure during and following spray treatments. Currently the only harmonised risk mitigation phrase aimed at reducing the exposure and hence the risk to bees is the following SPe8 from Annex V of 1999/45/EC1.

Dangerous to bees./To protect bees and other pollinating insects do not apply on flowering crops./Do not use where bees are actively foraging./Remove or cover beehives during application and for (state time) after treatment./Do not apply when flowering weeds are present./Remove weeds before flowering./Do not apply before (state time).

This phrase, or at least parts of it, is considered to cover risk mitigation for honey bees, bumble bees and solitary bees. For additional considerations about this risk phrase see Section 13.3.

13.3 Possible risk mitigation measures and associated phrases

Presented below are risk mitigation measures and associated phrases for the exposure scenarios considered in the exposure and risk assessment Chapters. Please note that no assessment of the risk to bees from honey dew is proposed in the current Guidance Document because the available information was not sufficient to produce a robust risk assessment scheme for this exposure route. However, recommendations for risk mitigation were included below which MSs could follow in case that there are concerns about risk to bees from exposure to honey dew.

13.3.1 Mitigation of the risk from spray applications

Scenario	Risk mitigation measures
Treated crop	SPe8 phrase
Treated weeds in the field	SPe8 phrase The SPe8 phrase includes a reference to removing weeds; therefore, it may be possible to use this phrase if a high risk to bees foraging treated weeds is predicted. However, possible adverse effects on biodiversity should be noted. Alternatively, the phrase "do not use where bees are actively foraging" also addresses this risk and permits flowering weeds to be left in place.
Field margin	If a risk is indicated from spray application on to a field margin, then it may be appropriate to consider the use of drift reducing measures, for example: *- Dangerous to bees./To protect bees and other pollinating insects, [specify risk mitigation measure, e.g. 90 % drift reducing spray nozzles, a buffer zone of x m, …] must be used.*
Adjacent crop	If a risk is predicted from spray application on to an adjacent crop, then it may be appropriate to consider the use of drift reducing measures, for example: *- Dangerous to bees./To protect bees and other pollinating insects, [specify risk mitigation measure, e.g. 90 % drift reducing spray nozzles, a buffer zone of x m, …] must be used.* *- Do not apply when the adjacent crops are flowering.*

Following crops	If a risk to succeeding non-permanent or following crops is predicted, it may be appropriate to propose a waiting period for bee-attractive succeeding crops and hence a phrase along the following lines may be used: *- Due to the risk to bees, bee-attractive crops should not be sown or planted within a period of [x] after [application/sowing/planting in the field].* It is not possible to mitigate the risk from permanent succeeding crops.

4513

4514 ### 13.3.2 Mitigation of the risk from applications of treated seeds or granules

4515 Risk mitigation measures from application of treated seeds are also covered in the draft
4516 Guidance document on treatment, placing on the market and use of treated seeds under
4517 Regulation (EC) No 1107/2009 (European Commission, 2012). The guidance document is
4518 currently undergoing revision, but might be considered in the future.

Scenario	Risk mitigation measures
Treated crop	SPe8 phrase: If a risk is predicted due to residues of the active substance and/or metabolites in nectar and pollen, it may be possible to mitigate the risk to honey bees by informing beekeepers that the crop has been grown from treated seed. This may be problematic owing to the foraging range of honey bees, location of colonies in relation to fields drilled with treated seed, etc.
Treated weeds in the field	Not relevant.
Field margin, adjacent crop	Risk from dust will occur in each sowing process from treated seeds or granules respectively. In case a standard seed quality has been considered in the risk assessment, it should be ensured that drilling of treated seed only occurs under certain conditions (e.g. with respect to wind conditions or the use of technical devices). An overview is given in Appendix II of the draft Guidance document on treatment, placing on the market and use of treated seeds under Regulation (EC) No 1107/2009 (European Commission, 2012). If a certain seed quality has been considered in the risk assessment (e.g. according to the proposed exposure quantification in the draft Guidance document for the authorisation of plant protection products for seed treatment – risk assessment) it would be necessary to certify the seed quality and even monitor the seed treatment process.
Following crops	If a high risk to succeeding or following crops is predicted, it may be appropriate to propose a waiting period for bee-attractive succeeding crops and hence a phrase along the following lines may be used: *- Due to the risk to bees, bee-attractive crops should not be sown or planted within a period of [x] after [application/sowing/planting in the field].* It is not possible to mitigate the risk from permanent succeeding crops.

4519

4520 ## 13.4 Developing risk mitigation measures

4521 If a high risk to bees cannot be addressed considering the currently available risk mitigation
4522 measures additional risk mitigation measures might need to be developed and proposed at
4523 first instance by the applicant. Therefore, if high risks are highlighted as part of the risk

assessment process, then it may be necessary to develop bespoke risk mitigation phrases to address the concern(s) highlighted.

When developing such measures, the following important factors have to be considered. The risk mitigation measure should be effective to address the identified concern. Ensure that all risk mitigation phrases are practicable and enforceable.

Always ensure that the risk mitigation phrase is seen by the relevant person. This is usually straightforward for spray formulations, where the risk mitigation can be stated on the product label. However, it is more complicated for treated seeds. For measures relevant to the sowing process of treated seed, the risk mitigation phrases should be on the bag with treated seed or accompanying document and not only on the seed treatment product label; see 1107/2009, Article 49.4.

For bee-attractive succeeding crops, these risk mitigation phrases should accompany plants that are grown from treated seed and then sold on to an end-user.

A SETAC publication titled "Mitigating the risks of plant protection products in the environment, MAgPIE" (Alix A, 2017) was published following the workshops held in 2013. The aim of the workshops was to produce a toolbox of risk mitigation measures which can be used in a quantitative risk assessment for bees. Even though the outcome of these workshops is not endorsed some examples of proposed mitigation measures are presented below. It is important to highlight that the proposal from the MAgPIE report should be considered as indicative.

In addition, the feasibility of the single phrases in context of agricultural and beekeeping practices was discussed.

- It may be appropriate to add the wording '…except as directed on [crop]' following '. crop plants when in flower,' where use is on several crops—but use on only some of these crops poses a risk to bees.
- There is uncertainty regarding the practicality of the phrase 'Remove or cover beehives during application and for [state time] after treatment', as it may result in an impact on honey bee colonies, for example from overheating. There are also practical difficulties as beekeepers may not live or be in the vicinity of their hives at the time of application. This risk mitigation measure is, of course, not relevant for bumble bees and solitary bees.
- The phrase 'Do not apply before (state time)' is potentially unclear, and hence clarification on the label would be required whether 'time' implies time of day, time of year or time in relation to crop development. It should also be noted bumble bees and solitary bees may forage at different times of the day and hence this phrase may protect one group but not the others.
- The recommendation 'remove weeds before flowering' is likely to have undesired side effects such as removing a source of nectar and pollen, which in turn may impact on honey bees, solitary bees and bumble bees. An analysis of monitoring studies (Alix A, 2017) was undertaken to describe the influence of farmland management practices. It was shown that removal of flowering weeds has an impact on bee species abundance due to gaps in foraging resources. Hence, in a wider perspective any reduction of food sources is considered to be a counterproductive risk mitigation measure for bees.

Based on the discussions the following revision of the SPe8 phrases was proposed.

- Dangerous to bees./To protect bees and other pollinating insects do not apply to crop plants when in flower./Do not use where bees are actively foraging./Remove or cover beehives during application and for (state time) after treatment./Do not apply when flowering weeds are present./ Do not apply before (state time)./Respect a flowering strip of [width to be specified] at [distance to be specified] of the treated field.
- Alert beekeepers prior to applying the product to allow adequate mitigation measures to be taken, and avoid bee colonies' exposure.

In addition to the SPe8 phrase new SPe phrases or modifications of current SPe phrases were discussed and proposed at the MAgPIE Workshop (Alix A, 2017) which address the risk to pollinators.

- To protect pollinators, respect an application rate of maximum (application rate to be specified)/do not apply this product more than (time period or frequency to be specified)/restrict applications to (dates or growth stages to be specified).
- To protect insects/pollinators and limit risks related to situations of runoff, respect an unsprayed non-cropped vegetated buffer zone of (distance to be specified) to the edge of the field which should consist of [wild flower mix/pollen and nectar mix] in order to provide the requested benefits.

Further, it was discussed to adapt the current SPe3 to reduce the exposure to pollinators in the untreated neighbouring areas (i.e. off-crop/off-field).

- To protect insects/pollinators from spray drift respect an unsprayed buffer zone of (distance to be specified) to the edge of the field. The edge of the field is either the edge of the crop or, in the presence of a margin strip, the edge of a margin strip.
- The buffer zone may be adjusted as a function of wind speed, wind direction, and temperature conditions based on available recommendations.
- The buffer zone may be reduced to (distance to be specified) if a combination of spray drift reduction technologies such as drift reducing nozzles, special equipment to reduce spray drift or directed spraying technique [is/are] used providing at least (% of drift reduction to be specified).

14 Conclusions

An updated risk assessment for honey bees, bumble bees and solitary bees is presented in the current document. The review of the EFSA (2013) has been performed in line with the ToRs of the mandate.

15 Recommendation

Placeholder: recommendations/need to further research will be developed]

16 References

Alix A, Brown C, Capri E, Goerlitz G, Golla B, Knauer K, Laabs V, Mackay N, Marchis A, Poulsen V, Alonso Prados E, Reinert W, Streloke M 2017. *Mitigating the Risks of Plant Protection Products in the Environment: MAgPIE*.

Baveco, J. M., Focks, A., Belgers, D., van der Steen, J. J., Boesten, J. J. & Roessink, I. 2016. An energetics-based honeybee nectar-foraging model used to assess the potential for landscape-level pesticide exposure dilution. PeerJ, 2016;4(e2293 10.7717/peerj.2293

Benfenati, Emilio, Como, Francesca, Manzo, Marco, Gadaleta, Domenico, Toropov, Andrey & Toropova, Alla 2017. Developing innovative in silico models with EFSA's OpenFoodTox database.

Camp, A. A., Batres, M. A., Williams, W. C. & Lehmann, D. M. 2020. Impact of Diflubenzuron on Bombus impatiens (Hymenoptera: Apidae) Microcolony Development. Environ Entomol, 2020;49(1):203-210 10.1093/ee/nvz150

Como, F., Carnesecchi, E., Volani, S., Dorne, J. L., Richardson, J., Bassan, A., Pavan, M. & Benfenati, E. 2017. Predicting acute contact toxicity of pesticides in honeybees (Apis mellifera) through a k-nearest neighbor model. Chemosphere, 2017;166(438-444 10.1016/j.chemosphere.2016.09.092

Corbet, Sarah A 2003. Nectar sugar content: estimating standing crop and secretion rate in the field. Journal of Apidologie, 2003;34(1):1-10

Dafni, Amots, Kevan, Peter G & Husband, Brian C 2005. *Practical pollination biology*.

Danforth, Bryan N, Minckley, Robert L, Neff, John L & Fawcett, Frances 2019. *The solitary bees: biology, evolution, conservation*, Princeton University Press.

Delaplane, K. S., Van Der Steen, J. & Guzman-Novoa, E. 2013. Standard methods for estimating strength parameters of Apis mellifera colonies. Journal of Apicultural Research, 2013;52(1):

Devillers, J., Pham-Delègue, M. H., Decourtye, A., Budzinski, H., Cluzeau, S. & Maurin, G. 2002. Structure-toxicity modeling of pesticides to honey bees. SAR and QSAR in Environmental Research, 2002;13(7-8):641-648 10.1080/1062936021000043391

Dimitrov, Sabcho, Dimitrova, Gergana, Pavlov, Todor, Dimitrova, Nadezhda, Patlewicz, Grace, Niemela, Jay & Mekenyan, Ovanes 2005. A stepwise approach for defining the applicability domain of SAR and QSAR models. Journal of chemical information and modeling, 2005;45(4):839-849 10.1021/ci0500381

Duan, Xiaodong, Wallis, David, Hatjina, Fani, Simon-Delso, Noa, Bruun Jensen, Annette & Topping, Christopher John 2022. ApisRAM Formal Model Description.

ECHA, (European Chemicals Agency), 2008. Guidance on information requirements and chemical safety assessment: Chapter R.6: QSARs and grouping of chemicals. Helsinki, Finland.

EFSA, (European Food Safety Authority) 2019. Outcome of the Pesticides Peer Review Meeting on general recurring issues in ecotoxicology. EFSA Supporting Publications, 2019;

EFSA, (European Food Safety Authority), Auteri, Domenica, Arce, Andres, Ingels, Brecht, Marchesi, Marco, Neri, Franco Maria, Rundlöf, Maj & Wassenberg, Jacoba 2022. Analysis of the evidence to support the definition of Specific Protection Goals for bumble bees and solitary bees.

EFSA, (European Food Safety Authority), Ippolito, Alessio , Aguila, Monica del , Aiassa, Elisa , Guajardo, Irene Muñoz , Neri, Franco Maria, Alvarez, Fernando, Mosbach-Schulz, Olaf & Szentes, Csaba 2020. Review of the evidence on bee background mortality.

EFSA, (European Food Safety Authority), Ippolito, Alessio, Focks, Andreas, Rundlöf, Maj, Arce, Andres, Marchesi, Marco, Neri, Franco Maria, Rortais, Agnès, Szentes, Csaba & Auteri, Domenica 2021. Analysis of background variability of honey bee colony size.

EFSA, (European Food Safety Authority) 2009. Risk Assessment for Birds and Mammals.

EFSA, (European Food Safety Authority) 2011. Submission of scientific peer-reviewed open literature for the approval of pesticide active substances under Regulation (EC) No 1107/2009.

EFSA, (European Food Safety Authority) 2013. EFSA Guidance Document on the risk assessment of plant protection products on bees (Apis mellifera, Bombus spp. and solitary bees). EFSA Journal, 2013;11(7):268 10.2903/j.efsa.2013.3295

EFSA, (European Food Safety Authority) 2014a. Guidance on the assessment of exposure of operators, workers, residents and bystanders in risk assessment for plant protection products.

EFSA, (European Food Safety Authority) 2014b. Guidance on expert knowledge elicitation in food and feed safety risk assessment. EFSA Journal, 2014b;12(6):3734

EFSA, (European Food Safety Authority) 2016. A mechanistic model to assess risks to honeybee colonies from exposure to pesticides under different scenarios of combined stressors and factors.

EFSA, (European Food Safety Authority) 2017. EFSA Guidance Document for predicting environmental concentrations of active substances of plant protection products and transformation products of these active substances in soil: This guidance published on 19 October 2017 replaces the earlier version published on 28 April 2015. EFSA Journal, 2017;15(10):e04982

EFSA, (European Food Safety Authority) 2018a. Peer review of the pesticide risk assessment for bees for the active substance clothianidin considering the uses as seed treatments and granules.
EFSA, (European Food Safety Authority) 2018b. Peer review of the pesticide risk assessment for bees for the active substance imidacloprid considering the uses as seed treatments and granules.
EFSA, (European Food Safety Authority) 2018c. Peer review of the pesticide risk assessment for bees for the active substance thiamethoxam considering the uses as seed treatments and granules.
EFSA, (European Food Safety Authority) 2018d. Evaluation of the data on clothianidin, imidacloprid and thiamethoxam for the updated risk assessment to bees for seed treatments and granules in the EU. EFSA Journal.
EFSA PPR Panel, (EFSA Panel on Plant Protection Products and their Residues) 2010. Scientific opinion on the development of SPG options for environmental risk assessment of pesticides, in particular in relation to the revision of the guidance documents on aquatic and terrestrial ecotoxicology (SANCO/3268/2001 and SANCO/10329/2002). EFSA Journal, 2010;8(1821
EFSA PPR Panel, (EFSA Panel on Plant Protection Products and their Residues) 2012. Scientific Opinion on the science behind the development of a risk assessment of Plant Protection Products on bees (Apis mellifera, Bombus spp. and solitary bees). EFSA Journal, 2012;10(5):2668
EFSA PPR Panel, (EFSA Panel on Plant Protection Products and their Residues) 2014. Scientific Opinion on good modelling practice in the context of mechanistic effect models for risk assessment of plant protection products.
EFSA PPR Panel, (EFSA Panel on Plant Protection Products and their Residues), Hernandez Jerez, Antonio, Adriaanse, Paulien, Berny, Philippe, Coja, Tamara, Duquesne, Sabine, Focks, Andreas, Marinovich, Marina, Millet, Maurice, Pelkonen, Olavi, Pieper, Silvia, Tiktak, Aaldrik, Topping, Christopher, Widenfalk, Anneli, Wilks, Martin, Wolterink, Gerrit, Rundlöf, Maj, Ippolito, Alessio, Linguadoca, Alberto, Martino, Laura, Panzarea, Martina, Terron, Andrea & Aldrich, Annette 2022a. Statement on the active substance flupyradifurone.
EFSA PPR Panel, (EFSA Panel on Plant Protection Products and their Residues), Hernandez Jerez, Antonio, Adriaanse, Paulien, Berny, Philippe, Coja, Tamara, Duquesne, Sabine, Focks, Andreas, Marinovich, Marina, Millet, Maurice, Pelkonen, Olavi, Pieper, Silvia, Tiktak, Aaldrik, Topping, Christopher, Widenfalk, Anneli, Wilks, Martin, Wolterink, Gerrit, Rundlöf, Maj, Ippolito, Alessio, Linguadoca, Alberto, Martino, Laura, Panzarea, Martina, Terron, Andrea & Aldrich, Annette 2022b. Statement on the active substance acetamiprid.
EFSA Scientific Committee 2016. Guidance to develop specific protection goals options for environmental risk assessment at EFSA, in relation to biodiversity and ecosystem services. EFSA Journal, 2016;14(6):e04499
EFSA Scientific Committee, More, Simon, Bampidis, Vasileios, Benford, Diane, Bragard, Claude, Halldorsson, Thorhallur, Hernández‐Jerez, Antonio, Bennekou, Susanne Hougaard, Koutsoumanis, Kostas & Machera, Kyriaki 2021. A systems‐based approach to the environmental risk assessment of multiple stressors in honey bees. EFSA Journal, 2021;19(5):e06607
EFSA Scientific Committee, Hardy, Anthony, Benford, Diane, Halldorsson, Thorhallur, Jeger, Michael John, Knutsen, Helle Katrine, More, Simon, Naegeli, Hanspeter, Noteborn, Hubert, Ockleford, Colin, Ricci, Antonia, Rychen, Guido, Schlatter, Josef R, Silano, Vittorio, Solecki, Roland, Turck, Dominique, Benfenati, Emilio, Chaudhry, Qasim Mohammad, Craig, Peter, Frampton, Geoff, Greiner, Matthias, Hart, Andrew, Hogstrand, Christer, Lambre, Claude, Luttik, Robert, Makowski, David, Siani, Alfonso, Wahlstroem, Helene, Aguilera, Jaime, Dorne, Jean-Lou, Fernandez Dumont, Antonio, Hempen, Michaela, Valtueña Martínez, Silvia, Martino, Laura, Smeraldi, Camilla, Terron, Andrea, Georgiadis, Nikolaos & Younes, Maged 2017. Guidance on the use of the weight of evidence approach in scientific assessments.
European Commission, 2012. Draft Guidance document on authorisation of plant protection products for seed treatment. European Commission document reference SANCO/10553/2012 rev. 2021.
European Commission, 2014. Assessing Potential for Movement of Active Substances and their Metabolites to Ground Water in the EU" Report of the FOCUS Ground Water Work Group. European Commission document reference Sanco/13144/2010 v. 3.
FOCUS, (Forum for the Co-ordination of Pesticide Fate Models and their Use) 2001. FOCUS surface water scenarios in the EU evaluation process under 91/414/EEC. Report of the FOCUS Working Group on Surface Water Scenarios. (updated by Generic Guidance for FOCUS surface water scenarios, v. 1.4, May 2015). EC Document Reference SANCO/4802/2001-rev. 2, 2001;245 pp

Foster, Robin L, Brunskill, Ameilia, Verdirame, David & O'Donnell, Sean 2004. Reproductive physiology, dominance interactions, and division of labour among bumble bee workers. Journal of Physiological Entomology, 2004;29(4):327-334

Gardner, Kathryn E, Foster, Robin L & O'Donnell, Sean 2007. Experimental analysis of worker division of labor in bumblebee nest thermoregulation (Bombus huntii, Hymenoptera: Apidae). Journal of Behavioral Ecology Sociobiology, 2007;61(5):783-792

Goulson, Dave, Peat, James, Stout, Jane C, Tucker, James, Darvill, Ben, Derwent, Lara C & Hughes, William OH 2002. Can alloethism in workers of the bumblebee, Bombus terrestris, be explained in terms of foraging efficiency? Journal of Animal Behaviour, 2002;64(1):123-130

Gradish, Angela E, Van der Steen, Jozef, Scott-Dupree, Cynthia D, Cabrera, Ana R, Cutler, G Christopher, Goulson, Dave, Klein, Olaf, Lehmann, David M, Lückmann, Johannes, O'Neill, Bridget, Raine, Nigel E, Sharma, Bibek & Thompson, Helen 2018. Comparison of Pesticide Exposure in Honey Bees (Hymenoptera: Apidae) and Bumble Bees (Hymenoptera: Apidae): Implications for Risk Assessments. Environmental Entomology, 2018;48(1):12-21 10.1093/ee/nvy168 %J Environmental Entomology

Hamadache, M., Benkortbi, O., Hanini, S. & Amrane, A. 2018. QSAR modeling in ecotoxicological risk assessment: application to the prediction of acute contact toxicity of pesticides on bees (Apis mellifera L.). Environ Sci Pollut Res Int, 2018;25(1):896-907 10.1007/s11356-017-0498-9

Hayward, A., Beadle, K., Singh, K. S., Exeler, N., Zaworra, M., Almanza, M. T., Nikolakis, A., Garside, C., Glaubitz, J., Bass, C. & Nauen, R. 2019. The leafcutter bee, Megachile rotundata, is more sensitive to N-cyanoamidine neonicotinoid and butenolide insecticides than other managed bees. Nat Ecol Evol, 2019;3(11):1521-1524 10.1038/s41559-019-1011-2

Heinrich, B 2004. *Bumblebee economics*, Harvard University Press.

Human, H., Brodschneider, R., Dietemann, V., Dively, G., Ellis, J. D., Forsgren, E., Fries, I., Hatjina, F., Hu, F. L., Jaffe, R., Jensen, A. B., Kohler, A., Magyar, J. P., Ozkyrym, A., Pirk, C. W. W., Rose, R., Strauss, U., Tanner, G., Tarpy, D. R., van der Steen, J. J. M., Vaudo, A., Vejsnaes, F., Wilde, J., Williams, G. R. & Zheng, H. Q. 2013. Miscellaneous standard methods for Apis mellifera research. Journal of Apicultural Research, 2013;52(4):55 10.3896/ibra.1.52.4.10

IPBES, 2016. The assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on pollinators, pollination and food production. SIMON G. POTTS, V. I.-F., HIEN T. NGO.

Jaworska, J., Nikolova-Jeliazkova, N. & Aldenberg, T. 2005. QSAR applicabilty domain estimation by projection of the training set descriptor space: a review. Altern Lab Anim, 2005;33(5):445-59 10.1177/026119290503300508

Kearns, Carol Ann & Inouye, David William 1993. *Techniques for pollination biologists*, University press of Colorado.

Knight, Mairi E, Martin, Andrew P, Bishop, Stephen, Osborne, Juliet L, Hale, Roddy J, Sanderson, Roy A & Goulson, Dave 2005. An interspecific comparison of foraging range and nest density of four bumblebee (Bombus) species. Journal of Molecular ecology, 2005;14(6):1811-1820

Kyriakopoulou, Katerina, Kandris, Ioannis, Pachiti, Irene, Kasiotis, Konstantinos M, Spyropoulou, Anastasia, Santourian, Anais, Kitromilidou, Stella, Pappa, Gerta & Glossioti, Maria 2017. Collection and analysis of pesticide residue data for pollen and nectar. EFSA Journal, 2017;14(10):

Last, G., Lewis, G. & Pap, G., 2019. Regulatory report on the occurrence of flowering weeds in agricultural fields. ERM.

Meixner, Marina D, Pinto, Maria Alice, Bouga, Maria, Kryger, Per, Ivanova, Evgeniya & Fuchs, Stefan 2013. Standard methods for characterising subspecies and ecotypes of Apis mellifera. Journal of Apicultural Research, 2013;52(4):1-28

Michener, Charles Duncan 2000. *The bees of the world*, JHU press.

More, Simon J, Auteri, Domenica, Rortais, Agnès & Pagani, Steve 2021. EFSA is working to protect bees and shape the future of environmental risk assessment.

Netzeva, T. I., Worth, A., Aldenberg, T., Benigni, R., Cronin, M. T., Gramatica, P., Jaworska, J. S., Kahn, S., Klopman, G., Marchant, C. A., Myatt, G., Nikolova-Jeliazkova, N., Patlewicz, G. Y., Perkins, R., Roberts, D., Schultz, T., Stanton, D. W., van de Sandt, J. J., Tong, W., Veith, G. & Yang, C. 2005. Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. The report and recommendations of ECVAM Workshop 52. Altern Lab Anim, 2005;33(2):155-73 10.1177/026119290503300209

Nieto, Anna, Roberts, Stuart PM, Kemp, James, Rasmont, Pierre, Kuhlmann, M, García Criado, Mariana, Biesmeijer, Jacobus, Bogusch, Pedr, Dathe, Holger H & De la Rúa, Pilar, 2014. European Red List of bees. Luxembourg.

Nikolova, Nina & Jaworska, Joanna 2003. Approaches to measure chemical similarity‐a review. Journal of QSAR Combinatorial Science, 2003;22(9 ‐ 10):1006-1026

NV, VITO 2021. Software tool for calculating the predicted environmental concentrations (PEC) of plant protection products (PPP) in soil for permanent and annual crops: Bug fixing & update report.

OECD, (Organisation for Economic Co-operation and Development) 1998a. *Test No. 213: Honeybees, Acute Oral Toxicity Test*, OECD Publishing.

OECD, (Organisation for Economic Co-operation and Development) 1998b. *Test No. 214: Honeybees, Acute Contact Toxicity Test*.

OECD, (Organisation for Economic Co-operation and Development) 2007. *Test No. 501: Metabolism in Crops*.

OECD, (Organisation for Economic Co-operation and Development) 2013. *Test No. 237: Honey Bee (Apis Mellifera) Larval Toxicity Test, Single Exposure*.

OECD, (Organisation for Economic Co-operation and Development) 2016. Guidance document on honey bee (Apis mellifera L.) Larval Toxicity Test, Repeated Exposure. Series on Testing and Assessment, No. 239, 2016;ENV/CBC/MONO(2016) 34(

OECD, (Organisation for Economic Co-operation and Development) 2017a. *Test No. 247: Bumblebee, Acute Oral Toxicity Test*.

OECD, (Organisation for Economic Co-operation and Development) 2017b. *Test No. 246: Bumblebee, Acute Contact Toxicity Test*.

OECD, (Organisation for Economic Co-operation and Development) 2017c. *Test No. 245: Honey Bee (Apis Mellifera L.), Chronic Oral Toxicity Test (10-Day Feeding)*.

OECD, (Organisation for Economic Co-operation and Development) 2021. Guidance document on honey bee (Apis mellifera L.) homing flight tests, using single oral exposure to sublethal doses of test chemicals. Series on Testing and Assessment, No. 332, 2021;ENV/CBC/MONO(2021) 7(

OEPP/EPPO, (European and Mediterranean Plant Protection Organization/Organisation européenne et méditerranéenne pour la protection des plantes) 2010. Efficacy evaluation of plant protection products: side ‐ effects on honey bees. PP 1/170 (4). OEPP/EPPO Bull., 2010;40(313-319

Ritz, C. 2010. Toward a unified approach to dose-response modeling in ecotoxicology. Environ Toxicol Chem, 2010;29(1):220-9 10.1002/etc.7

Roessink I, Hanewald N, Schneider C, Exeler N, Schnurr A, Molitor A-M, Soler E, Kimmel S, Molitor C, Smagghe G, Steen S van der. A method for a solitary bee (Osmia sp.) first tier acute contact and oral laboratory test: an update. Hazards of pesticides to bees - 13th international symposium of the ICP-PR Bee protection group, 2017 Valencia, Spain.

Roessink I, Hanewald N, Schneider C, Quambusch A, Exeler N, Cabrera AR, Molitor A-M, Tanzler V, Hodapp B, Albrecht M, Brandt A, Vinall S, Rathke A-K, Giffard H, Soler E, Schnurr A, Patnaude M, Couture A, Lehman D. Progress on the Osmia acute oral test - findings of the ICPPR Non-Apis subgroup solitary bee laboratory testing. Hazards of pesticides to bees - 14th international symposium of the ICP-PR Bee protection group, 2019 Bern, Switzerland.

Scheiner, R., Abramson, C. I., Brodschneider, R., Crailsheim, K., Farina, W. M., Fuchs, S., Grunewald, B., Hahshold, S., Karrer, M., Koeniger, G., Koeniger, N., Menzel, R., Mujagic, S., Radspieler, G., Schmickl, T., Schneider, C., Siegel, A. J., Szopek, M. & Thenius, R. 2013. Standard methods for behavioural studies of Apis mellifera. Journal of Apicultural Research, 2013;52(4):58 10.3896/ibra.1.52.4.04

Seeley, Thomas D 1982. Adaptive significance of the age polyethism schedule in honeybee colonies. Journal of Behavioral ecology sociobiology, 1982;11(4):287-293

Seeley, Thomas D 2009. *The wisdom of the hive: the social physiology of honey bee colonies*, Harvard University Press.

Sgolastra, Fabio, Hinarejos, Silvia, Pitts-Singer, Theresa L, Boyle, Natalie K, Joseph, Timothy, Lūckmann, Johannes, Raine, Nigel E, Singh, Rajwinder, Williams, Neal M & Bosch, Jordi 2019. Pesticide Exposure Assessment Paradigm for Solitary Bees. Environmental Entomology, 2019;48(1):22-35 10.1093/ee/nvy105 %J Environmental Entomology

Sinclair, C. J. & Boxall, A. B. 2003. Assessing the ecotoxicity of pesticide transformation products. Environ Sci Technol, 2003;37(20):4617-25 10.1021/es030038m

Singh, K. P., Gupta, S., Basant, N. & Mohan, D. 2014. QSTR modeling for qualitative and quantitative toxicity predictions of diverse chemical pesticides in honey bee for regulatory purposes. Chem Res Toxicol, 2014;27(9):1504-15 10.1021/tx500100m

Strange, James P, Garnery, Lionel & Sheppard, Walter S. 2007. Persistence of the Landes ecotype of Apis mellifera mellifera in southwest France: confirmation of a locally adaptive annual brood cycle trait. Journal of Apidologie, 2007;38(3):259-267

Toropov, Andrey A. & Benfenati, Emilio 2007. SMILES as an alternative to the graph in QSAR modelling of bee toxicity.

Venko, K., Drgan, V. & Novič, M. 2018. Classification models for identifying substances exhibiting acute contact toxicity in honeybees (Apis mellifera)$. SAR and QSAR in Environmental Research, 2018;29(9):743-754 10.1080/1062936X.2018.1513953

Wcislo, William T. & Cane, James H. 1996. Floral Resource Utilization by Solitary Bees (Hymenoptera: Apoidea) and Exploitation of Their Stored Foods by Natural Enemies. Annual Review of Entomology, 1996;41(1):257-286 10.1146/annurev.en.41.010196.001353

Wilson, E. O. & Holldobler, B. 2005. Eusociality: Origin and consequences. Proceedings of the National Academy of Sciences of the United States of America, 2005;102(38):13367-13371 10.1073/pnas.0505858102

Zurbuchen, Antonia, Landert, Lisa, Klaiber, Jeannine, Müller, Andreas, Hein, Silke & Dorn, Silvia 2010. Maximum foraging ranges in solitary bees: only few individuals have the capability to cover long foraging distances. Journal of Biological conservation, 2010;143(3):669-676

Zuur, Alain F, Ieno, Elena N, Walker, Neil J, Saveliev, Anatoly A & Smith, Graham M 2009. *Mixed effects models and extensions in ecology with R,* New York, NY, Springer.

17 Glossary and/or abbreviations and/or acronyms

[Placeholder: A full list of abbreviation and/or acronyms will be developed]

Appendices to the guidance document

Appendix A – List of crop attractiveness for pollen and nectar

This Appendix includes the revised list of crop for their attractiveness to bees for pollen and nectar. It is available as an excel spread sheet.

Appendix B – Parameters for contact and dietary exposure

This Appendix includes the parameters for contact and exposure estimation for honey bees, bumble bees and solitary bees. It is available as two separate excel spread sheets

Appendix C – Working examples for the lower tier risk assessment calculations for honey bees

C.1. Standard example honey bees

For an example calculation, it is assumed that for all four considered bee life stage and exposure type combinations a dose-response curve can be properly defined, i.e. well-defined $LD_{50,j}$ and $slope_j$ values can be determined for these four risk cases. Exposure doses are calculated as PEQ_j based on the definition of the relevant EREQ, here for the screening level.

The example is performed for the following GAP

Crop	BBCH	Application method	# Applications	Interval between applications	Application rate
Phacelia (Sugar content class 3)	40-69	Spray DW	2	7 days	200 g/ha

The GAP includes a pre-flowering and a flowering phase. As standard practice, the worst-case is assumed, i.e. that both applications are performed during flowering, which is well possible considering the rather short interval between applications and the flowering duration of Phacelia.

Considering that the application is performed directly on the flowering crop, the 'treated crop' scenario represents the worst-case, thus the risk assessment focus uniquely on this scenario.

Honey bees Screening level exposure [a]					
		Dietary			**Contact**
		Acute (da)	**Chronic (dc)**	**Larvae (dl)**	**Acute (ca)**
Hazard parameters: $LD_{50,j}$ [µg/bee] [b] and $slope_j$		$LD_{50,da}$ = 10.5 $slope_{da}$ = 1.84	$LD_{50,dc}$ = 6.75 $slope_{dc}$ = 1.67	$LD_{50,dl}$ = 4.75 $slope_{dl}$ = 2.24	$LD_{50,ac}$ = 12.4 $slope_{ac}$ = 2.23
Exposure (screening)	Factor B	6.4	6.2	7.2	-
	PEQ_j [µg/bee] [b]	PEQ_{da} = 2.56	PEQ_{dc} = 2.48	PEQ_{dl} = 2.88	PEQ_{ca} = 2.28
Step 1: Predicted individual level effect (PIE)		PIE_{da} = 6.9%	PIE_{dc} = 15.8%	PIE_{dl} = 24.6%	PIE_{ca} = 2.2%
Step 2: Predicted colony level effect (PCE)		PCE_{da} = 6.9%	PCE_{dc} = 15.8%	PCE_{dl} = 24.6%	PCE_{ca} = 2.2%
Step 3: combination of effects at colony level		$PE_{SPG} = 100 \cdot (1-(1-PCE_{da}/100) \cdot (1-PCE_{dc}/100) \cdot (1-PCE_{dl}/100) \cdot (1-PCEc_a/100))$ $= 100 \cdot (1-(1-0.069) \cdot (1-0.158) \cdot (1-0.246) \cdot (1-0.022))$ $=$ **42.2%**			

[a] *Note that for the contact exposure, there is no separate screening step. The calculated contact exposure therefore is equal to Tier-1.*

[b] *Units are mentioned for brevity as µg/bee, but they are in fact µg/bee/day for chronic and µg/bee/dev. Period for larvae*

In this example, the SPG is violated in the first step with an overall predicted effect at the colony level PE_{SPG} = 42.2%, using screening level PEQ_j values. Therefore, dietary exposure values are updated in an appropriate Tier-1 (see Chapter 5) and the overall predicted effect endpoint PE_{SPG} is calculated.

Honey bees Tier-1 exposure (Dietary and contact)					
		Dietary			**Contact**
		Acute (da)	**Chronic (dc)**	**Larvae (dl)**	**Acute (ca)**
Hazard parameters: $LD_{50,j}$ [µg/bee] [a] and $slope_j$		$LD_{50,da}$ = 10.5 $slope_{da}$ = 1.84	$LD_{50,dc}$ = 6.75 $slope_{dc}$ = 1.67	$LD_{50,dl}$ = 4.75 $slope_{dl}$ = 2.24	$LD_{50,ac}$ = 12.4 $slope_{ac}$ = 2.23
Exposure (Tier-1) [µg/bee] [a,b]	$SV_{ne,du}$	Forager = 3.9 Nurse = 2.1	Forager = 0.95 Nurse = 0.85	2.9	-
	$SV_{po,du}$	Forager = 0 Nurse = 3.5	Forager = 0 Nurse = 1.8	0.36	-
	PEQ_j	PEQ_{da} = 1.12	PEQ_{dc} = 0.53	PEQ_{dl} = 0.64	PEQ_{ca} = 2.28
Step 1: Predicted individual level effect (PIE)		PIE_{da} = 1.6%	PIE_{dc} = 1.4%	PIE_{dl} = 1.1%	PIE_{ca} = 2.2%

Step 2: Predicted colony level effect (PCE)	PCE_{da} = 1.6%	PCE_{dc} = 1.4%	PCE_{dl} = 1.1%	PCE_{ca} = 2.2%
Step 3: combination of effects at colony level	$PE_{SPG} = 100 \cdot (1-(1-PEC_{da}/100) \cdot (1-PCE_{dc}/100) \cdot (1-PCE_{dl}/100) \cdot (1-PCEc_{a}/100))$ $= 100 \cdot (1-(1-0.016) \cdot (1-0.014) \cdot (1-0.011) \cdot (1-0.022))$ = **6.2%**			

[a] *Units are mentioned for brevity as µg/bee, but they are in fact µg/bee/day for chronic and µg/bee/dev. Period for larvae*

[b] *whenever 2 sets of SVs were available for the specific risk case, the one between the two producing the worst-case exposure was used for PEQ_j calculations*

Using the more realistic model for dietary exposure estimation for Tier-1, resulted in calculated exposure quantity values for the dietary exposure that are lower. Consequently, an overall predicted effect at the colony level PE_{SPG} = 6.2%. Hence, using for both contact and for dietary exposure the Tier-1 exposure model to calculate the PEQ_j values, the protection goal for honey bees is not violated here.

In case Tier-1 RA would still breach the SPG, then further refinement of the PEQ can be done according to section 5.4.

Not for all substances the full set of hazard parameters ($LD_{50,j}$ and $slope_j$) will always be available for the risk cases. In case the derivation of a proper log-logistic dose-response curve is not possible because of limited data, in any case an LD50 value needs to be derived, and a default (conservative) slope value can be used as described in section 6.3. Leaving out one endpoint totally is not acceptable, unless an appropriate default PIE values is used.

C.2. Honey bees for a TRT substance

A second example calculation is performed for a substance that shows **time-reinforced toxicity (TRT)**. In this case, the dietary chronic predicted individual level effect is calculated based on the hazard parameters for the whole lifespan of a summer bee (27 days), and an exposure period of 27 days, instead of the 10 days as in the standard risk assessment. For the other risk cases and for the calculation of the PE_{SPG}, there are no differences from the standard risk assessment. The same use considered for the former example is also used here.

Honey bees – TRT - active period Tier -1 exposure (Dietary and contact)					
		Dietary			**Contact**
		Acute (da)	**Chronic (dc) – 27 days**	**Larvae (dl)**	**Acute (ca)**
Hazard parameters: $LD_{50,j}$ [µg/bee] [a] and $slope_j$		$LD_{50,da}$ = 10.5 $slope_{da}$ = 1.84	$LD_{50,dc}$ = 1.25 $slope_{dc}$ = 1.67	$LD_{50,dl}$ = 4.75 $slope_{dl}$ = 2.24	$LD_{50,ac}$ = 12.4 $slope_{ac}$ = 2.23
Exposure (Tier-1) [µg/bee] [a,b]	$SV_{ne,du}$	Forager = 3.9 Nurse = 2.1	0.39	2.9	-
	$SV_{po,du}$	Forager = 0 Nurse = 3.5	0.62	0.36	-
	PEQ_j	PEQ_{da} = 1.12	PEQ_{dc} = 0.20	PEQ_{dl} = 0.64	PEQ_{ca} = 2.28
Step 1: Predicted individual level effect (PIE)		PIE_{da} = 1.6%	PIE_{dc} = 4.6%	PIE_{dl} = 1.1%	PIE_{ca} = 2.2%
Step 2: Predicted colony level effect (PCE)		PCE_{da} = 1.6%	PCE_{dc} = 4.6%	PCE_{dl} = 1.1%	PCE_{ca} = 2.2%
Step 3: combination of effects at colony level		$PE_{SPG} = 100 \cdot (1-(1-PEC_{da}/100) \cdot (1-PCE_{dc}/100) \cdot (1-PCE_{dl}/100) \cdot (1-PCEc_{a}/100))$ $= 100 \cdot (1-(1-0.016) \cdot (1-0.046) \cdot (1-0.011) \cdot (1-0.022))$ = **9.2%**			

[a] *Units are mentioned for brevity as µg/bee, but they are in fact µg/bee/day for chronic and µg/bee/dev. Period for larvae*

[b] *whenever 2 sets of SVs were available for the specific risk case, the one between the two producing the worst-case exposure was used for PEQ_j calculations*

The overall predicted effect at the colony level PE_{SPG} is 9.2%, higher than what was predicted without consideration of TRT for the active period. Nevertheless, in this example, the SPG is not violated. The TRT characteristics is leading here to more sensitive hazard parameters, the LD_{50} after 27 days is here estimated assuming a Haber's exponent of -1.7, but also to lower exposure, since the relevant period is no longer 10 days but instead 27 days as a typical lifespan of a summer bee.

In addition to the risk assessment for the active period as presented above, an additional risk assessment for the winter bees also has to be performed for substances that show TRT. In this assessment, only chronic dietary exposure through consumption of honey is considered. In this case, the dietary chronic predicted individual level effect is calculated based on the hazard parameters for the whole lifespan of a winter bee (assumed equal to 182 days), and an exposure period of 182 days.

Honey bees – TRT - winter scenario Tier-1 exposure		
		Dietary
		Chronic (dc) – 182 days
Hazard parameters: $LD_{50,j}$ and $slope_j$		$LD_{50,dc}$ = 0.049 $slope_{dc}$ = 1.67
Exposure	SV	0.123
	PEQ_j	PEQ_{dc} = 0.025
Step 1: Predicted individual level effect (PIE)		PIE_{dc} = 24.0%
Step 2: Predicted colony level effect (PCE)		PCE_{dc} = 24.0%
Step 3: combination of effects at colony level		PE_{SPG} = **24.0%**

In this example, the SPG is violated with an overall predicted effect on the winter bees at the colony level PE_{SPG} = 24.0%. In case the active period assessment for this compound has not led to violation of the SPG, but the winter bee scenario is leading to a breach of the SPG.

Appendix D – Additional information for metabolite risk assessment

D.1. Introduction

This appendix proposes guidance for conducting the metabolite risk assessment for bees in the EU regulatory context. The main objectives of this Appendix are:

- To provide further information on available source for residue data in pollen and nectar and
- To provide further guidance on how to use non-test methods (e.g. QSAR).

D.2. Sources for residue data in pollen and nectar

Ideally the risk assessment for metabolites should be based on data from residues studies of the relevant matrices pollen and nectar. Therefore, there is a need to ensure that the relevant available information in the dossier is identified by the applicant and by the risk assessor.

Residues in pollen and nectar resulting from the use of the active substance are addressed by different studies. Each metabolite identified at or above 10% TRR and/or 0.01 mg eq/kg in these specific studies can be considered relevant.

Residue studies in (surrogate) crops in pollen and nectar

Residue are intended to provide information on the major degradation products and to measure the residues of the active substance and its degradation products in the relevant matrices pollen and nectar. The submission of a residue trial is no standard data requirement but is the most appropriate approach addressing the identification of relevant metabolites in pollen and nectar matrices. According to the data requirements on plant metabolism studies (Section 6 of Regulation (EC) No. 283/2013) residue trials should be performed according to the proposed critical GAP. The test conditions (maximum number of applications, shortest interval between application and maximum application rate) should be considered to identify the highest residues which may reasonably arise and should be representative of the realistic conditions of the critical GAP. The same assumptions can also be made for the metabolites formed in the matrices pollen and nectar. It may not be practical to test all crops in which the product is proposed to be used. As a worst-case approach the submission of a residue trial performed in a surrogate flowering crop (e.g. Phacelia, oilseed rape, or sunflowers) might be considered for the purpose of identification of relevant metabolites in pollen and nectar. Also, for such studies the proposed critical GAP considering number of applications, application rate and interval between the applications should be considered. For the identification of relevant metabolites and the indication of the % formation of the metabolite one residue trial per regulatory zone might be considered sufficient. On further guidance on residue trials see Annex B of the Guidance Document.

Plant metabolism study in primary and rotational crops

Metabolism studies are intended to elucidate the degradation pathway of the active substance, identify the metabolism and/or degradation product produced and provide an estimate of the total residues in the various parts of the plant. These studies are designed to identify metabolites at usually one point in time, typically not during flowering stage. For crops that can be grown in rotation, and when the DT_{90} of the active substance and/or relevant soil metabolites is above 100 days, studies should be performed to allow the determination of the nature and extent of potential residue accumulation in rotational crops from soil uptake, and of the magnitude of residues in rotational crops under realistic field conditions.

Residue level in pollen and bee products

To address flowering crops and to identify relevant metabolites in pollen and nectar, studies conducted to address the data requirement on residue levels in pollen and bee products for melliferous crops (Section 6.10.1 of Regulation (EC) No. 283/2013) may be available or by higher tier exposure and effect studies have been conducted to refine the risk for bees. However, studies containing information on residues in pollen and honey are only available in specific cases and consequently data on metabolites measured in pollen and nectar will be scarce. If relevant data is available it has to be remembered that while for higher tier residue studies the results might be used in a quantitative way, results derived from residue studies

for pollen and bee products should only be used in a qualitative way. Residue levels in bee products like honey and pollen (sampled in the hive) are diluted/concentrated and therefore not comparable to residue levels in pollen and nectar.

The risk assessment for metabolites in pollen and nectar relies upon the residues studies and therefore there is a need to ensure that the correct information available in the dossier is identified by the applicant and by the risk assessor. In particular, it is recommended that the following information, if available in the residue part of the dossier (Section 6 of Regulation (EC) No. 283/2013), is also made available or referred to in the ecotoxicology part. Further guidance on the use of residue data to support the ecotoxicological assessment is provided in the EFSA (2019), Appendix C.

General consideration for residue data

When data on the occurrence of metabolites in pollen and nectar are not available, then it is necessary to identify each metabolite in any plant metabolism study. Plant metabolism studies should be used to cover the risk both from bees foraging on the treated crop and other plants (weeds and adjacent crops). Rotational crop studies should be included for bare soil applications covering the succeeding crop scenario.

Concentration threshold for consideration of the residue

Only identified residues that are present individually at or above 10% TRR and/or 0.01 mg (parent) equivalent (eq)/kg are considered relevant. Usually in metabolism studies the values are normalised to the molecular weight of the parent. This is then expressed as "eq". For each identified metabolite, the maximum percentage of the metabolite (TRR) in any plant metabolism study in any matrix except roots should be reported.

Relevance of plant parts and choice of the appropriate sampling timepoints (in case of multiple sampling timepoints)

Residues in all parts of a plant, except roots, should be considered for the risk assessment. If, for a given plant part, information on the concentration of the metabolite is available for several sampling timepoints, the timepoint resulting in highest residue concentration should be considered in order to be protective. However, in a next step patterns and trends observed in residue studies in pollen and nectar or in plant metabolism studies can be taken into consideration in a case-by-case way depending on the intended GAP.

Information on the number of studies that were available for a particular crop and whether (the location of) the studies were spatially well distributed in EU shall be reported. It is assumed that all studies follow the GAP; exceptions should be noted. In case of exceptions (e.g. plant metabolism studies were conducted with higher application rates) the relevance of the metabolite should be considered in a case-by-case way.

D.3. Non-testing methods

D.3.1. Assessment based on parent toxicity data

Sinclair and Boxall (2003) investigated the toxicity of metabolites in relation to the parent compound of several PPPs (60 a.s. and 485 transformation products) and demonstrated that the majority (70 %) of transformation products have either a similar toxicity to the parent compound or are less toxic. However, a significant proportion (30 %) were more toxic than their parent compound and 4.2 % of transformation products were more than an order of magnitude more toxic.

Hence, as a first non-test strategy the metabolite risk assessment is based on the toxicity tests with the active substance assuming a 10-times higher toxicity compared to the parent. However, when it is clearly demonstrated that the metabolite is not expected to be more toxic than the parent, the metabolite can be assumed to be as toxic as the parent.

D.3.2. Identification of the toxophore

Substances that have a specific mode of action, such as pesticides, contain a structural feature or moiety that gives the toxic property. This structural feature is referred to as the toxophore, or toxophoric moiety. The substance causes toxicity through the interaction of its toxophore with a biomolecular site (e.g. receptor). Substances that are structurally similar could contain the same toxophore (or may yield a common toxophore upon metabolism) and may therefore have a common toxic effect.

For the assessment of the metabolite, it may be possible for the applicant to provide a reasoned case as to whether the molecule contains a toxophore or it has been lost following transformation. Toxophores for each of the major classes of PPPs have been identified by looking for substructural similarities within a pesticidal class by Sinclair (2009), which can be used to support argumentation. A number of ways have been identified to define domain of applicability, which may be used to decide whether or not toxophores are present ((Dimitrov et al., 2005); (Jaworska et al., 2005); (Netzeva et al., 2005); (Nikolova and Jaworska, 2003)). If it cannot be clearly shown that the toxophore is not present in the molecule, it should be assumed that the toxophore remains and that the molecule has a specific mode of action.

D.3.3. Generation of further data - (Q)SAR

The principles for assessing metabolites should in essence be the same as those for active substances, except for the acute contact toxicity. The acute contact study with metabolites is not required because exposure to metabolites will be predominantly via the oral route. As regards the issue of time-reinforced toxicity (TRT), if the tier 1 risk assessment for accumulative effects for the active substance indicates that this substance has accumulative effects, then it is assumed that the metabolite(s) will have as well. This is considered as being the worst case. In this situation, when the risk from the active substance is refined, it is important to consider the risk from the metabolite(s) as well.

However, in contrast to the active substance, data requirements for metabolites do not always have to be addressed by experimental studies. Applicants are invited to address the open questions by any other available information in support of a scientific and rational assessment, such as the use of non-testing methods like (Q)SAR/QSTR modelling as tools for deriving intrinsic properties of the metabolites of interest. The development and application of all kinds of non-testing methods is based on the similarity principle. The assumption is that similar compounds should have similar biological activities and the methods may therefore provide predictions that are reliable enough to substitute experimental data for several types of hazard related endpoints, for example, mortality. It has to be pointed out that a positive prediction made by the application of a non-test method for an effect may be accepted for use to avoid further testing while caution should be applied with negative predictions (i.e. lack of effect) since in most cases not all modes-of-action or mechanisms are covered by the existing non-test method.

When using (Q)SARs, it should be remembered that (Q)SARs are models and are therefore inevitably associated with a degree of uncertainty. This uncertainty is predominantly for two different reasons: (a) the inherent variability of the input data used to establish and validate

the (Q)SAR model and (b) the uncertainty resulting from the fact that a model can only be a partial representation of reality (in other words, it does not generally model all possible modes of action or mechanisms and hence does not represent all types of chemicals). It is noteworthy that these two types of uncertainty are related to the validation and the applicability domain of the (Q)SAR model respectively. Despite these uncertainties, it is also noted that a (Q)SAR is not only an empirical model, but that it is associated with (1) an underlying dataset used to establish and validate the model, (2) a description of the modelled endpoint, (3) the descriptors and the statistical methods used, (4) a characterisation of the applicability domain and (5) any appropriate mechanistic understanding of the model. As a representation of the training dataset for the model, it averages the uncertainty over all chemicals. Thus, if the model makes reliable predictions within its applicability domain, an individual model estimate will be more accurate than an individual measurement obtained by performing the relevant test.

Data provided by non-testing methods shall not be systematically used to fulfil the data requirements (Regulation (EU) No. 283/2013). However, there may be situations where non-testing methods can be used to address needs for information, rather than deriving new experimental data.

Right now, to the current knowledge, there are no QSAR models available (established/validated) to assess acute oral and/or chronic exposure of honey bees. However, ever since the early 90's efforts have been made to develop QSAR models for the acute contact toxicity of chemicals to bees (e.g (Benfenati et al., 2017), (Devillers et al., 2002), (Toropov and Benfenati, 2007), (Como et al., 2017), (Singh et al., 2014), (Venko et al., 2018) (Hamadache et al., 2018)). The expansion of public databases like e.g. EFSA's OpenFoodTox, US-EPA ECOTOX and Pesticide Properties Data-Base will hopefully provide (i) more measured biological activities for a set of molecules for the endpoint of interest (oral and/or chronic) and (ii) descriptions of the chemicals by means of their physicochemical properties, topological indices, and/or structural features, which will merit the development of statistical methods linking (i) and (ii) (QSAR model). Guidance on the validity of (Q)SAR models and reliability and adequacy of (Q)SAR predictions in general can be found in the ECHA report "Guidance on information requirements and chemical safety assessment Chapter R.6: (Q)SARs and grouping of chemicals" (ECHA, 2008).

Annexes to the guidance document

Annex A – Residue dissipation refinement

This Annex is available as standalone document, and it was developed jointly with the guidance document on Birds and Mammals. It is useful for risk assessment of terrestrial non-target organisms.

Annex B – Recommendations for residue trials to refine the exposure estimation

This Annex include the recommendations for performing residue trials useful to refine the exposure for honey bees, bumble bees and solitary bees at Tier 2.

Annex C – Recommendations for higher tier effect studies

This Annex include the recommendations for performing different type of higher tier effects studies to address the risk identified at lower tier risk assessment.