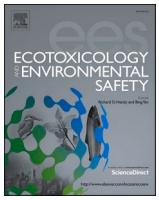
Ecotoxicology and Environmental Safety 285 (2024) 117062

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

journal homepage: www.elsevier.com/locate/ecoenv



Bee sensitivity derived from acute contact tests biased by standardised protocols?

Jan Baas [a,*], Timm Knautz [b], Annika Barme [b], Tatsuya Sekine [b], Marcel Jaklofsky [c], Dick Belgers [a], Nina Jansen [a], Marie-Claire Boerwinkel [a], Ivo Roessink [a]

[a] Wageningen University and Research, Droevendaalsesteeg 3, Wageningen 6708 PB, the Netherlands
[b] Ibacon GmbH, Arheilger Weg 17, Roßdorf 64380, Germany
[c] Wageningen University, department of Human and Animal Physiology, Wageningen 6708 WD, the Netherlands

ARTICLE INFO

Edited by Bing Yan

Keywords:
Bee testing
BeeGUTS
Acute contact test
Bee sensitivity

ABSTRACT

In an acute contact test with bees the compound of interest is dissolved in a carrier solvent (frequently acetone) and then a droplet of the solution is placed on the dorsal thorax of the bee. The volume of the droplet is standardised to 1 μL for honeybees and to 2 μL for bumblebees. In practice the same droplet volume is used for bees with very different sizes. In this research the effect of the droplet volume was evaluated with acute contact tests with dimethoate for the alfalfa leafcutter bee, the red mason bee, the honeybee and the bumblebee. The results were analysed with a ToxicoKinetic ToxicoDynamic (TKTD) model to separate kinetic from dynamic effects. This allows to compare the sensitivity of the bee based on the effect threshold and not on the time, species and test dependent LD_{50}s. The analysis of the test results indicates that the magnitude of the response of the bees increased with increasing droplet size. The results also showed that the manifestation of effects over time is slower for the red mason bee and the bumblebee compared to the honeybee and the alfalfa leafcutter bee. This implies that the result of a 2 day test with a fixed dosing volume results in different response for a bumblebee compared to the alfalfa leafcutter bee, not because of different sensitivities of the bees involved but due to the difference of relative dosed surface ratio. So comparing the sensitivity of bee species, based on standardised tests is biased and amplifies the sensitivity for the smaller bee species.

1. Introduction

Crop protection products are strictly regulated chemicals which require extensive testing to ensure that they will not pose risks to wildlife, plants and the environment. An important part of the overall testing for crop protection products is bee risk assessment. Bees play an important role in pollinating crops. There is a large variety of bee species; scientists have identified more than 20,000 species of bees, but most economically relevant crops are pollinated by a relatively small number of species (Kleijn et al., 2015). Of these, the honeybee is very important pollinator, even in regions where it is not native.

Standardised acute tests for honeybee testing date back to 1992 (EPPO, 1992; OECD 1998a, 1998b). More recently, in 2017, OECD guidelines were published for testing bumblebees (OECD, 2017). Solitary bees, like bees from the *Osmia* and *Megachile* family are also frequently tested, but there is no dedicated protocol for these species yet, but first steps have been taken for the development of standard tests for solitary bees. The importance of bee testing is evident but the large variety of bees requires insight in how to conduct and interpret tests in order to make accurate comparisons on the sensitivity of different bee species. Due to the easy handling and rearing most testing is carried out with honeybees (*Apis mellifera*). Therefore the bulk of available data is based on honeybee tests, which in turn raises the question on whether or not other bee species are protected if an environmental risk assessment is based on honeybees (Arena, 2014; Thompson, 2016). To date, comparisons on the sensitivity of bees (apis and non-apis) to chemicals are mainly based on the standardised acute contact tests, regardless of the size of the bee (Valdovinos-Núñez et al., 2009; Arena, 2014; Thompson, 2014; Uhl, 2016; Thompson, 2019).

The general setup of an acute contact test for honeybees is to administer a droplet of 1 μL, containing the compound of interest on the dorsal thorax of the adult worker bees, though other volumes can be used if justified (OECD 1998b). The compound of interest is usually dissolved in acetone or in water with addition of a surfactant like Triton

* Corresponding author.
E-mail address: jan.baas@wur.nl (J. Baas).

https://doi.org/10.1016/j.ecoenv.2024.117062
Received 10 June 2024; Received in revised form 13 September 2024; Accepted 14 September 2024
Available online 18 September 2024

X-100. The OECD guideline prescribes to use five doses in a geometric series. The tests are carried out in triplicate, with 10 bees per replicate. A toxic standard (usually dimethoate) is included in the test series to test if the sensitivity of the cohort is within a defined range. Mortality is recorded daily during at least 48 h and compared with control values. If the mortality rate is increasing between 24 and 48 h whilst control mortality remains at an accepted level, it is appropriate to extend the duration of the test to a maximum of 96 h. The result of the test is respectively the 48, 72 or 96 h LD_{50}, which is calculated from the experimental data, assuming a constant exposure concentration.

In practice solitary bees like the alfalfa leaf cutter bee and bees from the *Osmia* family are treated with the same droplet volume as the honeybee, so the size of the bee is not taken into account. Only the bumblebee test is has a standard droplet volume of 2 μL (but also here other droplet volumes may be used if justified (OECD, 2017)). This raises the question whether using a similar droplet volume for bees with different sizes (an alfalfa leafcutter bee for instance is about a quarter of the weight of a honeybee and about one tenth of the weight of a bumblebee) influences the outcome of the test. For smaller bee species a larger relative surface area is covered in a test, which has the potential to affect both the kinetics and the sensitivity of the bee. Still, both tests have a standard duration of 48 h and the droplet size used in a test is also similar.

As was previously shown by Baas and colleagues (Baas, 2022), the LD_{50} that is derived from acute tests is based on the initial dose and the assumption that this dose is constant over time. However as was shown by Zaworra and co-workers and Nauen and co-workers by washing the bees at different points in time (Nauen, 2015; Zaworra, 2019), the dose on the dorsal thorax of the bee is not constant during a test but declines over time. Further research has shown that the decline of the dose on the bee is bee-specific (Baas, 2024). In a test with honeybees app. 45 % of the initially applied dose if left after 48 h but for red mason bee this is app. 2 % (based on (Baas, 2024)). This implies that:

1. The LD_{50} based on the initial dose and assuming the dose is constant over the exposure duration (as is current practice) underestimates the sensitivity of a bee, since the actual dose is lower than the initial dose.
2. The LD_{50} derived from an acute contact test should not be used to compare the sensitivity of different species of bees as the decline of the dose over time is species specific.

It can be argued that the OECD guideline is meant to standardise the testing of bees aiming at a reliable method to generate normalized data for the regulatory risk assessment. It does not aim to represent the exposure of a bee under a field relevant exposure scenario like overspray (though similar processes might play a role) and details of the exposure are considered in higher-tier risk assessment. But the standard experimental framework as described in the OECD guideline is used to derive LD_{50}s which are used to compare the toxicity of different compounds and different species of bees. If LD_{50}s for one species of bees are compared this is not a real problem as the ranking stays the same, but when different species of bees are compared this is no longer the case as the test result is species-specific. In this research we will show that the sensitivity of smaller bees is consistently overestimate in such comparisons. This type of insight will (hopefully) contribute to the debate on comparing sensitivity of different bees, a debate recently revived by the revised guidance document on risk assessment for bees by EFSA (EFSA, 2023).

The recently published BeeGUTS paper (Baas, 2022) showed that the physiology of the species tested and the specifics of the test, such as the declining exposure concentration over time, can indeed be taken into account using a ToxicoKinetic ToxicoDynamic (TKTD) approach. Thus showing that the sensitivity of different test species can be assessed independently from any confounding factors originating from the actual test set-up used. The BeeGUTS model is a TKTD model that allows to interpret test data on how fast effects develop over time (the toxicokinetics) in combination with an assessment of the sensitivity of the bee (the toxicodynamics).

To investigate the impact of the usage of the standardized honeybee test guideline on the resulting LD_{50} values, dedicated acute contact tests were carried out with four different bee species that were all treated with different droplet volumes to see if kinetics or dynamics are indeed influenced by the test setup.

2. Materials and methods

2.1. General experimental setup

Test were carried out with 4 species of bees; the bees were chosen with the aim to cover a large size difference with some different characteristics. Typical size and weight characteristics are presented in Table 1.

With each bee species two independent tests were carried out:

1) measurement of the covered surface area of the insect with different droplet sizes
2) acute contact tests, also carried out with different droplet sizes

2.2. Origin of the bees and preparation for testing

Honeybees were acquired from queen-right colonies from the Team Environmental Risk Assessment of Wageningen University and Research. The colonies were housed at the Sinderhoeve experimental field station (Renkum, The Netherlands) which is situated in a rural area with only some minor extensive agriculture present. Honeybees were obtained from three separate colonies by collecting bees from the outer frames of the hive. Every colony populated one of the three replicates per treatment level used in the acute contact test. Further treatment and preparation of the bees was carried out according to OECD guidelines 213 and 214.

Bumblebees for the contact test originated from medium sized colonies, 10–12 weeks old, healthy, queen-right and having all brood stages. The colonies were obtained from a commercial bumblebee breeding company Koppert Deutschland GmbH, D-47638 Straelen. The bumblebees were collected from the upper, non-nest area of the plastic box under red light without the use of anaesthetics 1–2 days before application. Collection was carried out with plastic cages within one week after delivery. The bumblebees were collected from several colonies and were randomly allocated to the different treatment groups to avoid any colony effect within a treatment group. Application of dimethoate for bumblebees was conducted according to OECD guideline 246 (OECD, 2017b) after anesthetisation with carbon dioxide.

Red mason bees used in the 'droplet size' test, were obtained as cocoons from a local supplier (De Bijen Bestuivingstechniek, Nijmegen, The Netherlands). Cocoons were pre-selected and separated into males and females. Cocoons containing females are generally of larger size than those containing males. After arrival of the cocoons at the test

Table 1
Typical size and weight characteristics of the tested bee species.

Common name	Scientific name	Approximate weight (mg)	Approximate size *** (mm)
Honeybee	*Apis mellifera*	100*	11–13
Bumblebee	*Bombus terrestris*	300*	11 – 17
Red mason bee	*Osmia bicornis*	70** (average)	10 – 12 (female)
Alfalfa leafcutter bee	*Megachile rotundata*	30*	6–9

* Thompson et al. (Thompson, 2016)
** Uhl et al. (2016) (Uhl, 2016)
*** The bees of the world (Michener, 2007)

facility, they were stored in the refrigerator until use. Red mason bees for the contact test originated from cocoons obtained from a commercial solitary bee breeding company (WAB-Mauerbienenzucht, D-78467 Konstanz). Cocoons were pre-selected and separated into females and males.

For testing a suitable number of pre-selected female cocoons were placed in a flight cage. The cocoons were not manipulated in order to facilitate hatching and/or sexing of the bees (*i.e.* cutting the cocoons prior to hatching). Any males still present emerged earlier than the females and were removed from the cage daily. Emerged females were non-mated. They were stored in the refrigerator at 4 °C (± 4°C) with no food supply until enough solitary bees had been collected to start with the test. 18 h before the start of the application the emerged females were set in a flight cage and fed *ad libitum* before solitary bees were anaesthetized for application. Before start of the application the solitary bees were anaesthetized by chilling. CO_2 for anaesthesia can result in mortality for *Osmia* species and was therefore avoided. Solitary bees were anaesthetized by putting them for approx. 30 min in the refrigerator at 4 °C. During application the solitary bees were immobilized using an ice bath (solitary bees are placed in a plastic box which in turn is placed in iced water). The cold storage and chilling period were kept as short as possible. A single droplet of dimethoate in an appropriate carrier (tap water + 0.1 % v/v Triton X-100) was placed on the dorsal bee thorax using a calibrated pipette (Multipette©, Eppendorf).

Alfalfa leafcutter bees (*Megachile rotundata*) were obtained from a commercial supplier (North Star Seed Ltd, Canada). Cocoons were stored at 3°C in a climate cabinet and when needed for the experiment, incubated at 30°C according to the protocol from the supplier. Emergence of males and females peaked at approximately 18 and 24 after starting incubation, respectively. Tests were performed using females only.

2.3. Surface area test

The tests were carried out with a colouring agent (eosine) with droplet sizes of 1 μL; 2 μL, 4 μL and 6 μL in the laboratories of Wageningen University and Research. All tests were carried out with acetone or tap water containing 0.1 % triton X-100 as a carrier solvent. Eosine was kindly provided by the laboratory of Human and Animal Physiology of Wageningen University.

Each bee was given a droplet on the dorsal thorax of the bee according to the OECD protocol for acute contact tests with honeybees (OECD 1998b). For practical reasons not all bees were treated with all droplet sizes, as it would have been not feasible to treat a small leafcutter bee with 6 μL of liquid without the liquid falling off. See Table 2 for an overview.

After the droplet containing eosin was applied on the bee, the bees were immediately analysed under UV light. For this purpose the scanning microscope made a picture showing the droplet size via the luminescence of the eosin. The total surface area of the insect and the area covered by the droplet was determined by manual measurement using the freely available software package ImageJ.

2.4. Acute contact test

The tests were carried out according to the respective OECD guidelines for honey bees (OECD 1998b) and bumblebees (OECD, 2017) but with different droplet volumes; the dose (in μg/bee) was kept constant for the different droplet volumes for each species of bee. Deviations from the OECD guidelines for the alfalfa leaf cutter bee and the red mason bee are listed below:

Red mason bees and alfalfa leaf cutter bees were cooled for ~20 and 15 min resp. at 4 degrees before the dosing and the start of the experiment and were not anaesthetised with CO_2. The red mason bees were kept at 22 +/- 2 °C, at a relative humidity 60 %, Light-dark rhythm 16:8 h. The alfalfa leaf cutter bees were kept at 25 +/- 2 °C at a relative humidity of 60 % and a light-dark rhythm of 16:8 h.

All tests were carried out with dimethoate, honeybee and alfalfa leafcutter bee tests were carried with technical grade dimethoate of 99.8 % purity and acetone as a carrier solvent. For the alfalfa leafcutter bee an additional test was carried out with water containing 0.1 % of Triton X-100 but this led to a high control mortality and therefore the test was not further evaluated. The bumblebee and red mason bee tests were carried out with a formulated product (Danadim progress grade 417 g/L), using tap water containing 0.1 % Triton X-100 (Merck Millipore).

The bumblebee and red mason bee tests were carried out in the laboratories of IBACON (bumblebee according to OECD, 2017b, red mason bee according to current recommendations of the non-*Apis* ring test group (ICPPR non-Apis group, 2016–2023)), the honeybee and alfalfa leafcutter bee tests were carried out in the laboratories of Wageningen University and Research.

Mortality was recorded on each day for a total of 96 h, in line with the OECD guidelines. Additional observation on lethal effects were carried out 4 h after the start of the experiment for the bumblebee (OECD, 2017) and the red mason bee. Additional observations on mortality were carried out at 4, 8, 12 and 32 h for the tests with the honeybee and the alfalfa leafcutter bee.

For practical reasons not all tests were carried out with all droplet sizes. Higher droplet volumes make no sense for the alfalfa leaf cutter bee and the standard droplet size for bumblebees is 2 μL which already leads to a low relative surface area covered. An overview of the tests is presented in Table 3. The bumblebee tests were carried out in duplicate.

2.5. Data interpretation of the acute contact tests

The test results were interpreted with the BeeGUTS model (Baas, 2022). This is a dedicated TKTD model for the interpretation of bee tests, based on the General Unified Threshold model for Survival (GUTS) modelling framework. A TKTD model for data-interpretation is different from the standard LD_{50} data analysis in that it takes the complete survival matrix as input for the model and not just the effects at a specific point in time and disregarding all other timepoints with observations on effects. The development of effects over time is followed and this contains kinetic information and information on the effect threshold. The latter is by definition the LD_0 at infinite exposure time. This effect threshold is important as this is a measure for the intrinsic toxicity of a compound/species combination. Another important advantage of TKTD data interpretation is that the exposure concentration does not need to be constant. Any exposure scenario (constant exposure, pulsed exposure, first order decline (as in the acute bee tests)) can be used as input for the

Table 2
Overview of the droplet size tests.

Name	Droplet size			
	1 μL	2 μL	4 μL	6 μL
Apis mellifera	*	*	*	
Bombus terrestris		*	*	*
Osmia bicornis	*	*	*	
Megachile rotundata	*	*		

Table 3
Overview of the acute contact tests that were carried out with different droplet sizes.

Name	Droplet size			
	1 μL	2 μL	4 μL	5 μL
Apis mellifera	*	*	*	
Bombus terrestris		*		*
Osmia bicornis		*		*
Megachile rotundata	*	*		

model.

The GUTS modelling framework unifies two independent models with different assumptions on the death mechanism: Stochastic Death (SD) and Individual Tolerance (IT). None of the models are considered superior and in practice both models are used and generally the most conservative result is used for further risk assessment. When internal and external concentrations (over time) are available, these can be used for the data interpretation, this is then called the full GUTS model. When the only external concentrations are available, formally then the reduced GUTS approach is used, as in this research. Both models use three parameters to describe the development of toxic effects over time. Details of GUTS approach and the underlying assumptions are described in great detail elsewhere (Jager, 2011; Ashauer, 2016; EFSA, 2018; Jager, 2018).

A more elaborate description of the BeeGUTS model with the parameters to calculate the exposure scenarios for the different bee species can be found in the Supporting Information (SI).

The result of the data analysis with the BeeGUTS model are three parameters that together describe the whole process of the development of toxic effects. In addition control mortality is estimated as a separate parameter:

- The effect threshold; as a measure for the intrinsic toxicity of the bee
- The dominant rate constant; as a measure for the kinetics
- The killing rate in the Stochastic Death model (a measure for the toxicity of the compound once the effect threshold is exceeded or the spread in the effect thresholds in the Individual Tolerance model
- Control mortality, which was estimated from the complete dataset, in line with the recommendation from Plantade and coworkers (Plantade, 2023)

The main advantage of an interpretation with a TKTD model is that this results in test-independent parameters, which allows for a separation of effects of the droplet size on the intrinsic sensitivity (by comparing the effect threshold) and the kinetics (by comparing the other parameter values).

3. Results

3.1. Results surface area test

In Fig. 1, an example of the pictures of the eosin test with honeybees and bumblebees is presented.

In Fig. 2, the covered surface area in mm^2 for the different tests is shown, both for acetone and for water containing 0.1 % of Triton X-100.

In Fig. 2, the surface of area covered for the different droplet sizes using acetone or water containing 0.1 % Triton X-100 as solvent or wetting agent in mm2 and as % of the whole bee and of the thorax size for the different droplet volumes is shown.

3.2. Results of the Acute Contact tests

Control mortality for bumblebees and honeybees in the tests carried out with a droplet volume of 1 and 2 μL droplet is <10 % at the end of the 4 d (96 hr) test and so in compliance with the OECD test protocols. Control mortality for honeybees in the test with a 4 μL droplet volume is slightly higher at the end of the 4 d test: 13 % (the 3 d (72 hr) control mortality was 10 %). This might be caused by the higher droplet size itself but this was not considered to be a reason to exclude the result at 4 days, with a TKTD data analysis control mortality is estimated from the complete dataset and not only on the controls, which gives a better representation of control mortality (Plantade, 2023). Control mortality in the Osmia test was also <10 % and control mortality in one of the alfalfa leaf cutter bee tests was between 17 % and 15 % at the end of the 3 d (72 h) test. Since the survival matrix was consistent, both over time and concentration, this gave no reason to reject any of the tests.

Dimethoate is frequently used as a positive control in bee testing and the OECD guidelines present requirements for the 2 d LD_{50}s derived from the test. The 2 d (48 h) LD_{50} value for dimethoate calculated according to the OECD protocol (assuming a constant exposure concentration) for honeybees carried out with a 1 μL droplet volume equals 0.13 (0.11–0.15) μg/bee and 4.9 (4.3–5.6) μg/bee for bumblebees (test 1) carried out with the 2 μL droplet volume. These are both in the required range given in the respective protocols, implying that the test results can be accepted. The 2 d (48 h) LD_{50} for the 2 and 4 μL droplet volume for honeybees were 0.12 (0.10–0.14) and 0.094 (0.097–0.12)

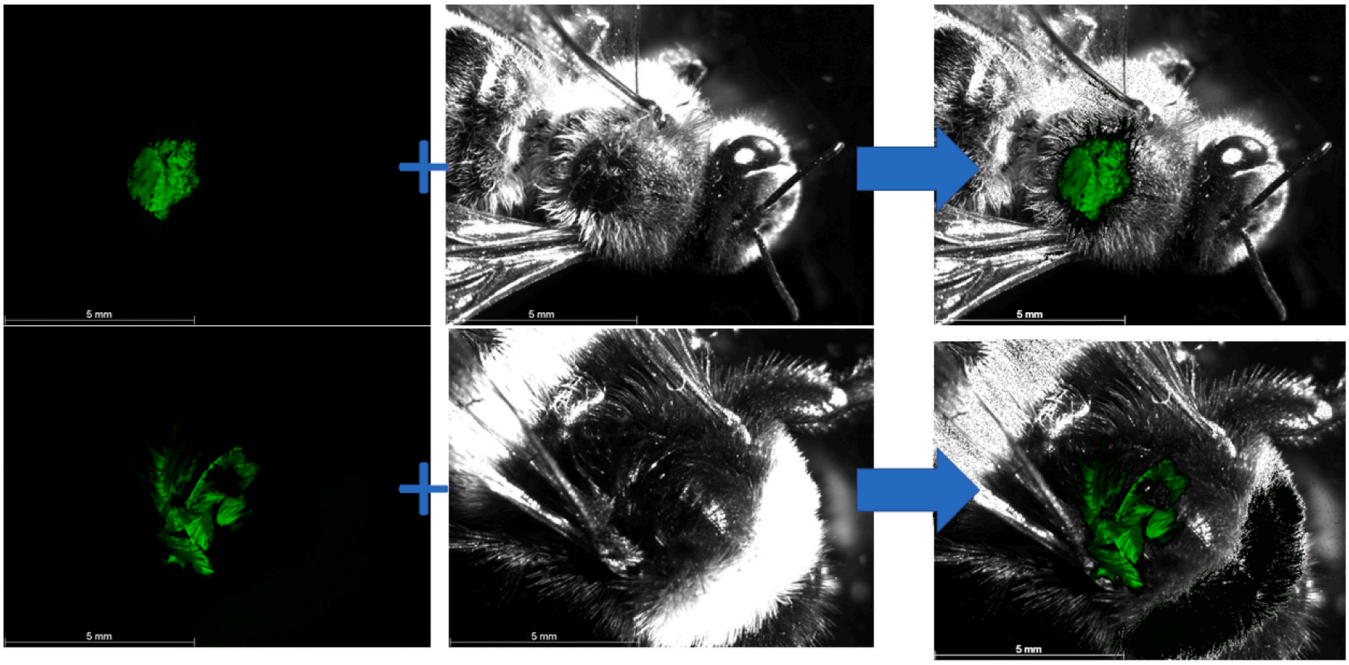


Fig. 1. Example of eosin stained droplet on a honeybee (top row) and bumblebee (bottom row).

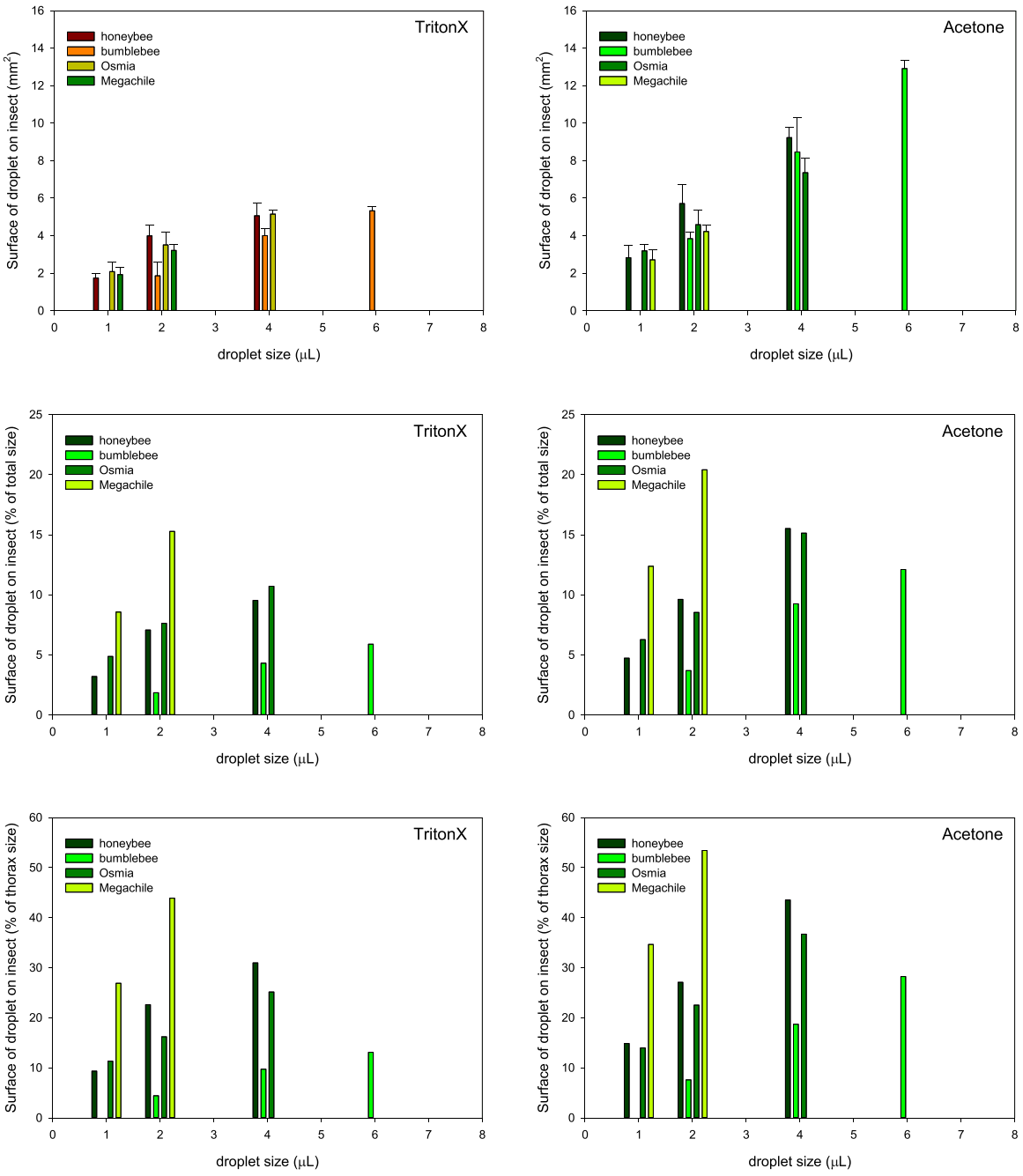


Fig. 2. Top panels: surface of covered area in mm^2 for the different droplet sizes using acetone or water containing 0.1 % Triton X-100 as solvent or wetting agent. Middle panels: percentage of total body size covered for the different droplet sizes. Lower panels: percentage of thorax surface area covered using either Triton X100 (left panels) or Acetone (right panels) as wetting agent or solvent.

µg/bee resp. The bumblebee test (test 1) carried out with a 5 µL droplet volume gave a 2 day (48 h) LD_{50} of 5.0 (4.4–6.4) µg/bee. An example of the survival matrix for honeybees is presented in Table 4.

The TKTD parameter estimates and the Normalised Root Mean Squared Error (NRMSE) values (a measure for the average deviation from the calculated vs the observed mortality) for the different tests are listed in Table 5 for the SD model. The results of the SD model and the IT model gave a comparable interpretation of the data, therefore only the results of the SD model are presented in tables in the main text. A more elaborate description of the results, including the goodness of fit values and the results of the IT model are presented in the SI.

All tests have survival matrices that are consistent, both over time and over concentration. The BeeGUTS model generally gives excellent fits of the data (R^2 values range between 0.92 and 0.98; NRMSE values range between 7 % and 25 %) so well within the range to accept the fits as outlined by the EFSA guideline on TKTD modelling (EFSA, 2018).

The parameters derived from the tests were used to calculate the LD_{50} values at 2 days (48 h) and 4 days (96 h), these are listed in Table 6. Note that these values are calculated based on the parameter values listed in Table 5, taking into account the decline of the concentration on the bee during a test. These values are therefore different from the values calculated according to the OECD protocol, where it is (wrongly) assumed that the exposure concentration is constant.

Table 4
Survival matrix for the honeybee test conducted with a droplet volume of 1 μL.

Time (hour)	Dose (μg/bee)					
	0 (control)	0.04	0.08	0.16	0.32	0.64
	Nr of survivors					
0	30	29	30	30	30	30
4	30	29	30	29	24	16
8	30	29	29	20	4	0
12	30	29	29	15	0	0
24	30	29	29	6	0	0
32	30	29	29	6	0	0
48	30	29	29	5	0	0
72	30	28	27	5	0	0
96	28	28	26	3	0	0

Table 5
Parameter values of the different tests when interpreted with the SD model.

Test	Volume (μL)	k_d (d^{-1})	m_w (μg/bee)	b_w (1/(μg/bee.d^{-1}))	NRMSE (%)
A. mellifera	1	6.1 (4.5–8.4)	0.086 (0.070–0.10)	43 (30−65)	8.6
A. mellifera	2	4.4 (3.4–5.9)	0.080 (0.065–0.093)	54 (37−79)	10.8
A. mellifera	4	3.9 (2.5–5.6)	0.056 (0.050–0.060)	42 (29−69)	14.5
B. terrestris, test 1	2	1.8 (1.2 – 3.0)	1.2 (0.80 – 1.5)	0.28 (0.20 – 0.39)	7.8
B. terrestris, test 2	2	3.0 (1.9 – 5.4)	2.2 (1.4 – 3.0)	0.28 (0.18 – 0.44)	10.5
B. terrestris, test 1	5	4.1 (2.6 – 7.2)	1.5 (1.0 – 1.8)	0.24 (0.17 – 0.33)	8.1
B. terrestris, test 2	5	3.7 (2.3 – 6.8)	2.2 (1.6 – 3.1)	0.24 (0.17 – 0.38)	6.8
O. bicornis	2	< 0.33	0.019 (9^E−4 – 0.076)	16 (4.9 – 1200)	7.3
O. bicornis	5	0.34 (0.12 – 0.62)	0.059 (0.018 – 0.11)	5.0 (2.7 – 18)	11.1
M. rotundata	1	7.6 (5.0–14)	0.072 (0.059–0.084)	23 (14−36)	9.3
M. rotundata	2	15 (7.0–65)	0.045 (0.021–0.054)	22 (12−31)	8.4

4. Discussion

4.1. Surface area test

The droplet size test shows that when using acetone as a carrier solvent a larger surface area of the bee is covered, both in absolute as in relative terms. Apparently, acetone enhances the spreading of the toxicant over the 'waxy' surface of the bees more than water containing 0.1 % Triton X-100.

The relative surface area covered was linearly related to the applied droplet size and inversely proportional to the size of the bee. When water containing 0.1 % Triton X-100 is used as a carrier solvent, a droplet size of 2 μL leads to a covered relative surface area of around 2 % for *B. terrestris*, about 7 % for *O. bicornis*, about 8 % for *A. mellifera* and about 15 % for *M. rotundata*. With acetone these covered relative areas comprise approximately 4, 9, 10 and 20 %, respectively.

Probably the covered area of the insect thorax is the most important factor for the uptake of the pesticide over the cuticle of the insect (with the same solvent). With a droplet volume of 1 μL in a test with honeybees, around 15 % of the area of the thorax is covered. But for the alfalfa leafcutter bee this area increases to approximately 35 % more than doubling the relative surface area impacted for the honeybee. The bumblebee in the standard OECD approach with a droplet size of 2 μL, receives a coverage of around 7 % of the surface area of the thorax. In this case, less than halve the area of the thorax is covered compared to the honeybee. The red mason bee is comparable in size to the honeybee and as a result the surface area covered is also comparable to that of the honeybee.

Table 6
Calculated 2 d and 4 d LD_{50} values based on the TKTD parameters for the different tests.

Test	Droplet Volume (μL)	Carrier solvent	Calculated 2 d LD_{50} (μg/bee)	Calculated 4 d LD_{50} (μg/bee)
A. mellifera	1	Acetone	0.097 (0.083–0.11)	0.091 (0.075–0.11)
A. mellifera	2	Acetone	0.090 (0.076–0.10)	0.084 (0.070–0.097)
A. mellifera	4	Acetone	0.068 (0.063–0.074)	0.061 (0.055–0.066)
B. terrestris, test 1	2	Triton X100	3.2 (2.8–3.7)	2.0 (1.7–2.3)
B. terrestris, test 2	2	Triton X100	4.0 (3.4–4.6)	3.0 (2.4–3.7)
B. terrestris, test 1	5	Triton X100	3.3 (2.9–3.8)	2.3 (1.9–2.6)
B. terrestris, test 2	5	Triton X100	4.1 (3.6–4.9)	3.1 (2.5–3.8)
O. bicornis	2	Triton X100	0.41 (0.34–0.48)	0.15 (0.12–0.21)
O. bicornis	5	Triton X100	0.45 (0.37–0.55)	0.19 (0.16–0.25)
M. rotundata	1	Acetone	0.090 (0.081–0.11)	0.081 (0.070–0.094)
M. rotundata	2	Acetone	0.063 (0.046–0.070)	0.054 (0.034–0.061)

If the results of acute contact tests need to be compared for different species of bees, covering a similar relative surface is the starting point. However, then there are practical limitations in reducing the droplet volume, therefore it is recommended to use a droplet volume of 1 μL for the alfalfa leafcutter bee, 2 μL for the honeybees and bees of similar size and of 5 μL for the bumblebees. For smaller bees than the alfalfa leaf cutter bee smaller droplet sizes are recommended as long as the solubility of the compound of interest is sufficient.

4.2. Acute contact tests

The tests with the red mason bee and the bumblebees were carried out without the additional observations on effects early in the test. The lack of additional observations on effects is reflected in the larger confidence intervals of the TKTD parameter values. For both species, the confidence intervals for the parameter values overlap and therefore differences in effects for the different droplet sizes cannot be derived from the data. The test with the bumblebees was carried out in duplo and here the differences in parameter values between the duplo's are comparable to the differences between the tests with different droplet sizes. For both the bumblebee and the red mason bee, the confidence intervals in the parameter values are too wide to find an effect of droplet size on the parameter values.

The tests with the honeybees and the alfalfa leaf cutter bees, were carried out with a dedicated protocol with additional observations on effects early in the test. The survival matrix shows that the main changes in the observed effects take place in the first part of the test, see Table 4. The main effects occur in the first 24 h of the test. This implies that the most important information on the parameter values of a TKTD model is in the first phase of the test, which is lacking with the standard test protocol, with observations on effects at 4, 24 and 48 h. Similar observations were done for the other droplet sizes and the alfalfa leafcutter bees (see SI).

From the honeybee and alfalfa leafcutter bee tests parameter values can be derived with relatively narrow confidence intervals, see Table 4, and therefore provide the best information on the effects of droplet size on the TKTD parameter values. As a result, the data generated with the dedicated protocol for the honeybee and the alfalfa leaf cutter bee were used for further evaluation of the effect of droplet size on the parameter values.

4.3. Effect of droplet size on the sensitivity of bees

The test with honey bees and alfalfa leaf cutter bees was carried out with acetone as a vehicle. The bumblebee test and the red mason bee tests were carried out with triton X-100 as a vehicle. This might have had an effect in itself so the comparisons of the different test results carried out with different vehicles must be carried out with some reserve. However the data do not suggest a significant influence of the vehicle.

Fig. 3 shows the development of effects over time for the honeybee and also shows the excellent goodness of fit of the model for the honeybees (similar figures for the other bees are presented in the SI). The differences in observed survival between the tests with the different droplet sizes are rather subtle and are best visible in the second and third treatment.

Despite the rather subtle differences in the observations on effects, both the results for the tests with honeybees and the alfalfa leafcutter bees both show that the sensitivity of the bees (expressed as the effect threshold) is affected by the volume of the droplet used in the acute test, see Fig. 4.

Also the calculated LD_{50} values indicate a higher sensitivity with increasing droplet size for the honeybee and the alfalfa leafcutter bee, see Table 5. For honeybees the increase in sensitivity from a treatment with a 1 μL droplet to a 4 μL droplet is about 25 %. For the alfalfa leaf cutter bee the increase in sensitivity in going from a 1 μL droplet to a 2 μL droplet is also around 25 %. Though the confidence intervals of the parameter values for the bumblebee and red mason bee tests are too large to directly draw this conclusion, the observations on effects also suggest an increased sensitivity with increased droplet size. The survival matrix shows that after 4 h of exposure for bumblebees a 10 % effect is observed in the highest exposure concentration in the 5 μL experiment but not in the 2 μL experiment. In addition, there is also a slight effect (3 %) in the highest exposure concentration in one of the 5 μL experiments but not in the 2 μL experiments.

It is also interesting to see that the difference between the 48 hr LD_{50} and the 96 h LD_{50} is rather small (< 10 %) for the honeybee and the alfalfa leafcutter bee (~10–15 %), but larger for the bumblebee (~25–40 %) and the red mason bee (~65 %). Or in other words the incipient LD_{50} is almost reached in the honeybee and alfalfa leafcutter bee tests with dimethoate in 48 h, but the bumblebee test and the red mason bee test need more time to reach the incipient LD_{50} for dimethoate. With the bumblebee test, this is most likely caused by the sheer size of the bee but for the red mason bee probably by its different, more bristle like, type of hair on its cuticle (Roquer-Beni, 2020).

The modelling results with the SD model suggest that for the bumblebee the slower development of effects over time is not caused by differences in the dominant rate constant but by the dynamics. The tests for honeybees, alfalfa leafcutter bees and the red mason bee all indicate a b_w value > 5 $(\mu g/bee.d^{-1})^{-1}$, whereas for bumblebees this value is significantly lower; around 0.28 $(\mu g/bee.d^{-1})^{-1}$. A (relatively) high value for b_w (with comparable values for the dominant rate constant)

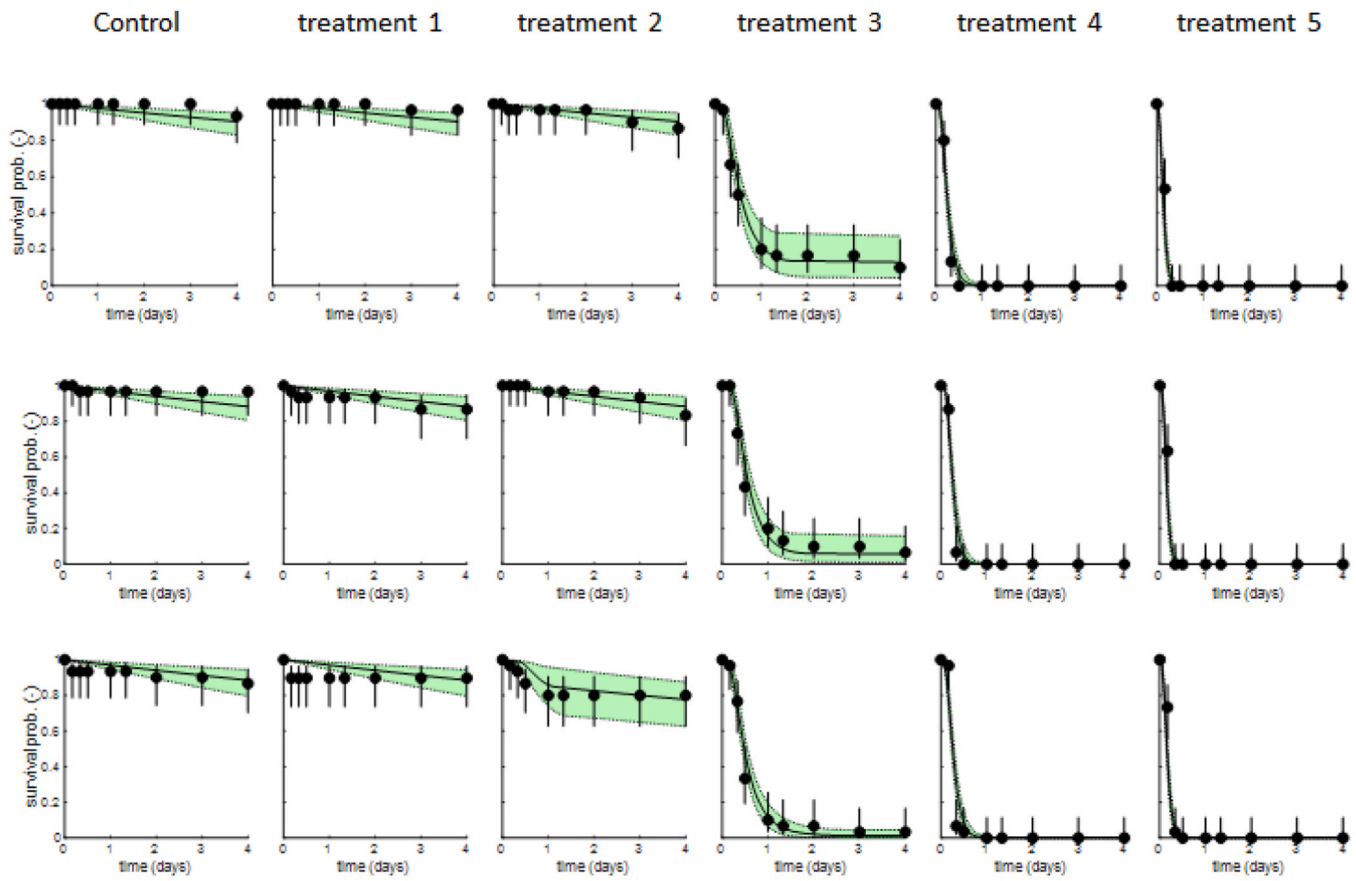


Fig. 3. Result of the honeybee tests. Top row, test with 1 μL droplet, second row test with 2 μL droplet, bottom row test with 4 μL droplet. All tests were carried out with one control and 5 different exposure concentrations. The dots represent the observed survival, the lines represent the calibrated model result and the green area is the 95 % confidence interval.

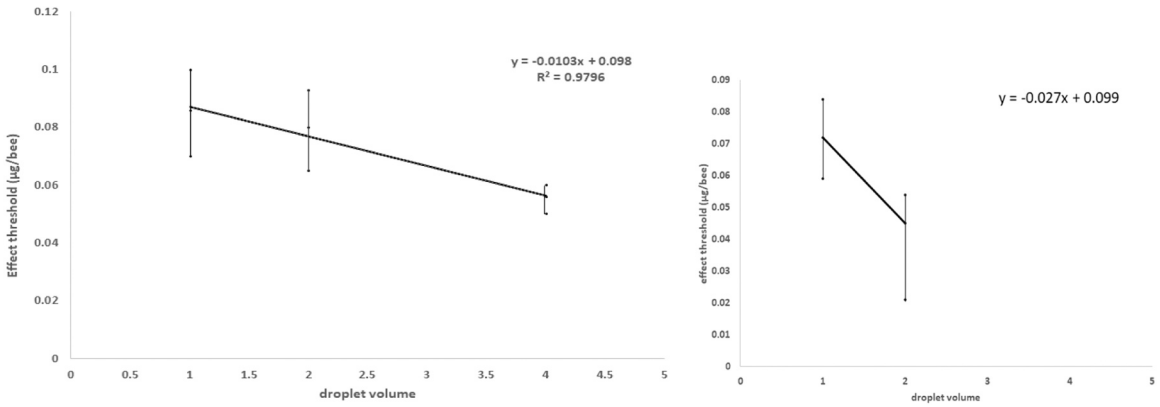


Fig. 4. Droplet volume vs effect threshold for honeybees (first panel) and alfalfa leafcutter bees (second panel).

implies that once the effect threshold is exceeded, mortality occurs relatively fast. In the extreme case where b_w is infinite, the organism dies immediately after the effect threshold is exceeded. So bumblebees are not only less sensitive for dimethoate than honeybees (reflected in a lower effect threshold for honeybees) but once the threshold concentration is exceeded the effects also take longer to develop. For the red mason bee the parameter values indicate that here it is the slower uptake kinetics (the K_d value is $< 0.3\ d^{-1}$, where this value is $> 1.8\ d^{-1}$ for other bee species) that causes the slower developments of effects over time. The results of the IT model show a lower value for the dominant rate constant for the bumblebee and the red mason bee compared to the honeybee and the alfalfa leafcutter bee. So with this model the slower development of the effects over time is reflected in the dominant rate constant.

Though the SD and IT models differ in their parameter values, the models do not differ in their interpretation of the data: time goes slower for bumblebees and red mason bees in the acute contact test with dimethoate compared to the honeybee and the alfalfa leafcutter bee. Or in other words 48 h of exposure for a bumblebee does not equal 48 h of exposure for a honeybee. The clock ticks faster for a honeybee and alfalfa leaf cutter bee and therefore effects have further developed falsely indicating a higher sensitivity.

This implies that a comparison of the sensitivity of a bee, based on its LD_{50} derived from a standard test must be carried out with care as this LD_{50} does not reflect the actual sensitivity of a bee. This is underlined by the observation that tests with the alfalfa leafcutter bee frequently show effects 4 h after the start of the exposure, where it is rare to already observe effects for honeybees or bumblebees 4 h after the start of an exposure.

5. Conclusions

Comparing the sensitivity of different species of bees based on the acute contact test should take into account:

- The exposure pattern of the test, because this is species dependent
- The exposure time as clock ticks faster for smaller species of bees
- The dominant rate constant and the effect threshold depend on the relative covered surface area

Comparing the toxicity of different compounds within one species based on the LD_{50}, that is based on the initial concentration and the assumption that the dose is constant over time, underestimates the actual toxicity of the compound but the ranking will still be valid.

To overcome these issues, it is recommended to compare the sensitivity of bees, based on the time- and test independent effect threshold, as this is a better proxy for the sensitivity of the bee. However, to obtain the best estimates of the effect threshold additional time points early in the test are required.

In addition, the droplet size should be adapted for different species of bees to start the test with a comparable covered relative surface area of the bees in a test. Droplet size advice: 1 µL for smaller bees like the alfalfa leaf cutter bee, 2 µL for medium sized bees like the honeybee and 5 µL for larger bees like the bumblebee.

CRediT authorship contribution statement

Tatsuya Sekine: Writing – review & editing, Validation, Investigation. **Marcel Jaklofsky:** Validation, Investigation, Data curation. **Dick Belgers:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Data curation. **Nina Jansen:** Validation, Investigation, Data curation. **Marie-Claire Boerwinkel:** Validation, Investigation, Data curation. **Ivo Roessink:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization. **Jan Baas:** Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Timm Knautz:** Writing – review & editing, Visualization, Validation, Supervision, Investigation, Data curation. **Annika Barme:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Investigation, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

All (raw) data are available in the supporting information.

Acknowledgements

Grant from the Dutch Ministry of Economic Affairs: BO43-102.01-007, Vanessa Roeben for help in creating the graphic abstract, Henk Schipper for assistance with the fluorescent analysis.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ecoenv.2024.117062.

References

Arena, M., Sgolastra, F., 2014. A meta-analysis comparing the sensitivity of bees to pesticides. Ecotoxicology 23, 324–334. https://doi.org/10.1007/s10646-014-1190-1.

Ashauer, R., Albert, C., Augustine, S., Cedergreen, N., Charles, S., Ducrot, V., Focks, A., Gabsi, F., Gergs, A., Goussen, B., Jager, T., Kramer, N.I., Nyman, A.-M., Poulsen, V., Reichenberger, S., Schäfer, R.B., Van den Brink, P.J., Veltman, K., Vogel, S., Zimmer, E.I., Preuss, T.G., 2016. Modelling survival: exposure pattern, species sensitivity and uncertainty. Sci. Rep. 6, 29178 https://doi.org/10.1038/srep29178.

Baas, J., Goussen, B., Miles, M., Preuss, T.G., Roessink, I., 2022. BeeGUTS—a toxicokinetic–toxicodynamic model for the interpretation and integration of acute and chronic honey bee tests. Environ. Toxicol. Chem. 41, 2193–2201. https://doi.org/10.1002/etc.5423.

Baas, J., Goussen, B., Taenzler, V., Roeben, V., Miles, M., Preuss, T.G., van den Berg, S., Roessink, I., 2024. Comparing sensitivity of different bee species to pesticides: A TKTD modeling approach. Environ. Toxicol. Chem. 43, 1431–1441. https://doi.org/10.1002/etc.5871.

EFSA, Adriaanse, P., Arce, A., Focks, A., Ingels, B., Jölli, D., Lambin, S., Rundlöf, M., Süßenbach, D., Del Aguila, M., Ercolano, V., Ferilli, F., Ippolito, A., Szentes, C., Neri, F.M., Padovani, L., Rortais, A., Wassenberg, J., Auteri, D., 2023. Revised guidance on the risk assessment of plant protection products on bees (Apis mellifera, Bombus spp. and solitary bees). e07989. DOI EFSA J. 21. https://doi.org/10.2903/j.efsa.2023.7989.

EFSA, Products Panel oP.P., Residues t, Ockleford, C., Adriaanse, P., Berny, P., Brock, T., Duquesne, S., Grilli, S., Hernandez-Jerez, A.F., Bennekou, S.H., Klein, M., Kuhl, T., Laskowski, R., Machera, K., Pelkonen, O., Pieper, S., Smith, R.H., Stemmer, M., Sundh, I., Tiktak, A., Topping, C.J., Wolterink, G., Cedergreen, N., Charles, S., Focks, A., Reed, M., Arena, M., Ippolito, A., Byers, H., Teodorovic, I., 2018. Scientific Opinion on the state of the art of Toxicokinetic/Toxicodynamic (TKTD) effect models for regulatory risk assessment of pesticides for aquatic organisms. e05377. DOI EFSA J. 16. https://doi.org/10.2903/j.efsa.2018.5377.

EPPO, 1992. Guideline on test methods for evaluation the side-effects of plant protection products on honeybees. In: Organisation, E.P.P. (Ed.), *170*. Bulletin OEPP/EPPO Bulletin, 22, pp. 203–215.

Jager, T., Albert, C., Preuss, T.G., Ashauer, R., 2011. General unified threshold model of survival - a toxicokinetic-toxicodynamic framework for ecotoxicology. Environ. Sci. Technol. 45, 2529–2540. https://doi.org/10.1021/es103092a.

Jager, T., Ashauer, R., 2018. How to evaluate the quality of toxicokinetic-toxicodynamic models in the context of environmental risk assessment. Integr. Environ. Assess. Manag. 14, 604–614. https://doi.org/10.1002/ieam.2026.

Michener C.D. 2007. The bees of the world. 2nd edition. John Hopkins University press, Baltimore.

Nauen, R., Jeschke, P., Velten, R., Beck, M.E., Ebbinghaus-Kintscher, U., Thielert, W., Wölfel, K., Haas, M., Kunz, K., Raupach, G., 2015. Flupyradifurone: a brief profile of a new butenolide insecticide. Pest Manag. Sci. 71, 850–862. https://doi.org/10.1002/ps.3932.

OECD. 1998a. Test No. 213: Honeybees, Acute Oral Toxicity Test.

OECD. 1998b. Test No. 214: Honeybees, Acute Contact Toxicity Test.

OECD. 2017. Test no. 246: OECD guideline for the testing of chemicals Bumblebee, acute contact toxicity test.

Plantade, J., Baudrot, V., Charles, S., 2023. hb or not hb - when and why accounting for background mortality in toxicological survival models matters? MethodsX 10, 102114. https://doi.org/10.1016/j.mex.2023.102114.

Roquer-Beni, L., Rodrigo, A., Arnan, X., Klein, A.-M., Fornoff, F., Boreux, V., Bosch, J., 2020. A novel method to measure hairiness in bees and other insect pollinators. Ecol. Evol. 10, 2979–2990. https://doi.org/10.1002/ece3.6112.

Thompson, H., 2016. Extrapolation of acute toxicity across bee species. Integr. Environ. Assess. Manag. 12, 622–626. https://doi.org/10.1002/ieam.1737.

Thompson, H.M., Fryday, S.L., Harkin, S., Milner, S., 2014. Potential impacts of synergism in honeybees (Apis mellifera) of exposure to neonicotinoids and sprayed fungicides in crops. Apidologie 45, 545–553. https://doi.org/10.1007/s13592-014-0273-6.

Thompson, H.M., Pamminger, T., 2019. Are honeybees suitable surrogates for use in pesticide risk assessment for non-Apis bees? Pest Manag. Sci. 75, 2549–2557. https://doi.org/10.1002/ps.5494.

Uhl, P., Franke, L.A., Rehberg, C., Wollmann, C., Stahlschmidt, P., Jeker, L., Brühl, C.A., 2016. Interspecific sensitivity of bees towards dimethoate and implications for environmental risk assessment. Sci. Rep. Vol 6.

Valdovinos-Núñez, G., Quezada-Euan, J.J., Ancona-Xiu, P., Moo-Valle, H., Carmona, A., Ruiz, E., 2009. Comparative toxicity of pesticides to stingless bees (Hymenoptera: Apidae: Meliponini). J. Econ. Entomol. 102, 1737–1742. https://doi.org/10.1603/029.102.0502.

Zaworra, M., Koehler, H., Schneider, J., Lagojda, A., Nauen, R., 2019. Pharmacokinetics of three neonicotinoid insecticides upon contact exposure in the Western Honey Bee, Apis mellifera. Chem. Res. Toxicol. 32, 35–37. https://doi.org/10.1021/acs.chemrestox.8b00315.