Integrated Environmental Assessment and Management — Volume 00, Number 00—pp. 1–11
Received: 6 February 2024 | Revised: 3 July 2024 | Accepted: 8 July 2024

Health & Ecological Risk Assessment

Honey bee larval toxicity study designs: Applicability of the current study protocols and endpoints as a predictor of pesticide hazard for pollinators

Daiana A. De Souza,[1,2] *Max Feken,*[2,3] *Hudson V. V. Tomé,*[2,4] *and Daniel R. Schmehl*[2,5]
[1]*BASF Corporation, Raleigh, North Carolina, USA*
[2]*Pollinator Research Task Force (PRTF), USA*
[3]*Syngenta Crop Protection, LLC., Greensboro, North Carolina, USA*
[4]*FMC Corporation, Newark, Delaware, USA*
[5]*Bayer Crop Science LP, Chesterfield, Missouri, USA*

Abstract

The assessment of pesticide risks to bees in North America currently relies in part on Tier 1 honey bee laboratory toxicity studies to support the registration and registration review processes for crop protection chemicals. For immature stages, the studies follow two standardized test designs recommended by the Organization for Economic Cooperation (OECD), evaluating acute (seven-day single-dose, TG OECD 237) and chronic (22-day repeated-dose, GD OECD 239) toxicity in bee larvae. In this article, we aim to evaluate the current approach for generating and interpreting honey bee larval toxicity data, enhancing pesticide risk assessment for pollinators. First, by considering that the repeated-dose larval study covers all stages of honey bee brood development up to adult emergence, we compared endpoints (larval LD/ED50 and LC/EC50 values) from seven-day acute exposure studies with the 22-day chronic exposure studies. Our goal was to identify the study design offering greater sensitivity in assessing pesticide toxicity to immature bees. Our second objective involved analyzing available weight data from emerged adults and comparing it to survival endpoints (e.g., NOEL and LD50) to determine if the weight after adult emergence would accurately represent a sensitive indicator of pesticide effects on developing honey bees. Our analysis determined that the use of a single 22-day chronic exposure study adequately covers all immature stages and that the toxicity values based on cumulative dose are more accurate and representative measures of exposure for immature bees than using endpoints based on estimated daily doses. Furthermore, our analysis suggests that measuring the weight of emerged adults was a more sensitive indicator than mortality of treatment-related effects in 22% of the compounds included in our analysis. Here we also discuss the importance of standardized protocols for proper collection of weight after emergence and the need for further discussion on the relevance of this parameter at risk assessment scheme. *Integr Environ Assess Manag* 2024;00:1–11. © 2024 SETAC

KEYWORDS: Emergence weight; Lethal dose; Pollinators; Risk assessment; Toxicity in bee larvae

BACKGROUND

The decline of pollinators around the globe has been enhancing the overall concern regarding their safety, when facing multiple environmental stressors, including pesticide exposure (Johnson, 2014; Siviter et al., 2023). Considering these issues from a regulatory standpoint, global agencies such as the USEPA, PMRA (Health Canada Pest Management Regulatory Agency), and EFSA (European Food Safety Agency) include protection goals for pollinators within the environmental risk assessment (ERA) framework for pesticide registration. Pesticides have multiple chemical classes with strict regulatory policies that require an in-depth ERA prior to receiving regulatory approvals for product use. In the United States, under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), pesticide use cannot result in unreasonable risk to human health and the environment (Bruce et al., 2023). The current ERA framework relies upon a tiered approach for assessing the risk of pesticides on nontarget organisms, including pollinators. The initial core data requirements are from laboratory-based studies that support a screening-level risk assessment that includes worst-case exposure estimates (i.e., including maximum label application rates and generic residue data) to estimate the RQ (risk quotient). When risks are not excluded at the lower-tier screening-level assessment, additional exposure and effects studies are conducted that simulate more realistic and complex semi-field and field scenarios that better represent real-world environmental risks (USEPA, 2014).

This article contains online-only Supporting Information.
Address correspondence to daiana.almeida-de-souza@basf.com
Published on wileyonlinelibrary.com/journal/ieam.

The honey bee (*Apis mellifera* L.) is the surrogate test species for assessing risk to terrestrial nontarget arthropods following the current USEPA ERA framework, although data availability on other species is also considered when available (USEPA, 2016). The honey bee laboratory-based studies are designed to assess the acute and chronic toxicity of a pesticide to both adult and larvae bees. The laboratory studies are conducted in accord with the Organization for Economic Cooperation and Development (OECD) and/or USEPA Office of Chemical Safety and Pollution Prevention (OCSPP) standardized test protocols and guidance documents, which aim to quantify effects at the individual level and include measurements for mortality and sublethal effects (USEPA, 2014). Honey bee adults may be exposed to pesticides via contact and dietary exposure. Similarly, honey bee larvae may be exposed to pesticide residues, yet are confined to individual cells and may be exposed to pesticides via oral and contact exposure through the diet (the larvae are partially submerged in the diet during development) as well as contact exposure via contaminated wax (Böhme et al., 2019; Milone et al., 2021; Sanchez-Bayo & Goka, 2014).

There are currently two standardized test designs to assess pesticide effects on larvae, a single exposure study with a seven-day duration following the OECD 237 test guideline, commonly referred to as acute study design (OECD, 2013) (Figure 1A), and a repeat-exposure study with a 22-day duration following OECD 239 test guidance, mostly referred as chronic study design (OECD, 2016) (Figure 1B). The seven-day "acute" larval study guideline was initially developed when technical limitations prevented a longer study duration to adult emergence due to a high rate of control mortality and inconsistent results (Aupinel et al., 2009). The seven-day acute study only assesses effects during the larval life stage. In the acute larval assay, first instar larvae are transferred from healthy colonies to grafting cells on Day 1 and a single dose of the test chemical is administered to the larvae with the diet on Day 4. Mortalities are recorded daily from Day 5 to Day 7 of the tests and the 72 h LD_{50} for larvae based on the cumulative mortality on Day 7 (OECD, 2013).

The advancement of in vitro larval rearing methods (Clark, 2017; Crailsheim et al., 2013; Schmehl et al., 2016) has enabled the development of the OECD 239 guidance document that includes four days of "chronic" exposure to the pesticide and has a test duration of 22 days (i.e., through adult emergence). Similar to the acute larval assay, first instar larvae are transferred from healthy colonies to grafting cells (Day 1) but instead, they are treated with a diet administered on Days 3 through 6, at a constant concentration in the diet equivalent to increasing test chemical doses per larva per day.

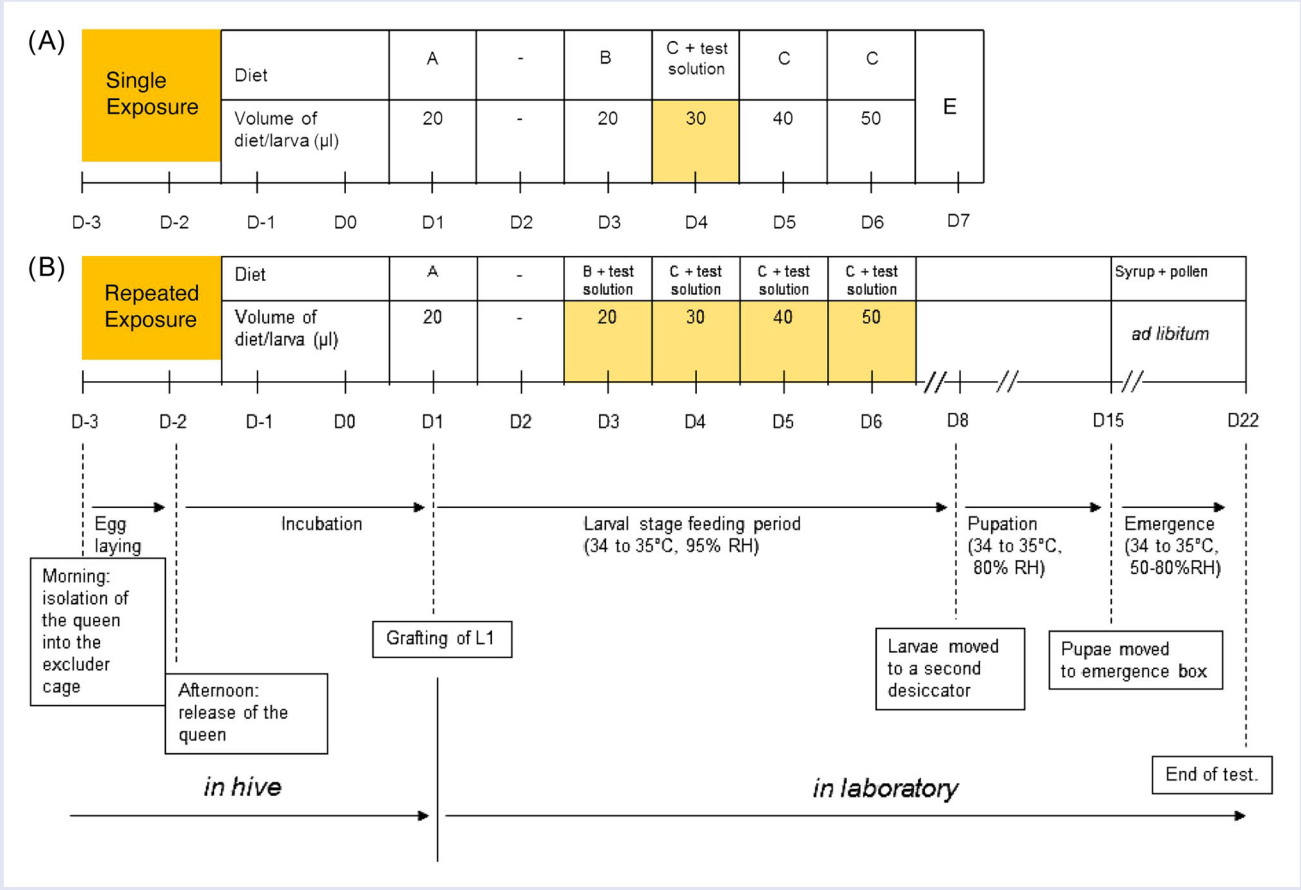

FIGURE 1 Schematic representation of important steps in the larval acute study design (A) and chronic toxistudy design; (B) Toxicity test. D, day; RH, relative humidity. *Source*: OECD (2013, 2016) modified

The 22-day chronic larval study measures effects during the larval and pupal phases up through adult emergence, with daily observations between Days 4 and 8 for larvae, Day 15 for pupae, and Day 22 for adult emergence (i.e., when the bee has completed its development to the adult stage) (OECD, 2016; Figure 1B). The chronic study design is focused on determining the No Observed Effect Concentration/Cumulative Dose (NOEC/NOED) based on Day 22 adult emergence and, if data allow, EC50/ED50 (or any ECx/EDx) on Day 22 (adult emergence) although these endpoints can also be determined at other observation periods (e.g., Day 8 for larvae). Some regulatory authorities, including the USEPA and IBAMA (Brazilian Institute of Environmental and Renewable Natural Resources), require both the seven-day acute exposure study and 22-day chronic exposure study as core data requirements, yet it is unclear whether the seven-day acute study would provide additional information to cooperate risk assessment when a longer 22-day chronic-exposure study is available. Our first study objective was to evaluate and compare endpoints (larval LD/ED50 and LC/EC50 values) from available seven-day acute (single dose) exposure studies with the 22-day chronic (repeat dose) exposure studies to determine if one study design results in lower endpoints (i.e., greater sensitivity) and can fully inform pesticide risks to immature bees.

The 22-day chronic exposure study design is predominantly used to quantify the cumulative No Observed Effect Dose/Concentration (NOED/NOEC) or the effective dose/concentration for 50% of the bees (ED50/EC50) based upon survival; however, the OECD Guidance Document (GD) 239 guidance permits for other observations to be recorded qualitatively, including adverse effects after emergence (OECD, 2016). Although the adult body weight after emergence measurement is not explicitly mentioned in OECD 239 GD as are other sublethal parameters (Maus et al., 2022), the USEPA has requested, in cases when the agency reviewed proposed protocols, the inclusion of weight after adult emergence to inform the growth assessment endpoint when determining the risk of a given pesticide to bees. The collection of data on weight after adult emergence has been generated in academic studies (Dai, Jack, Mortensen, Bloomquist, et al., 2018; Dai, Jack, Mortensen, Bustamante, et al., 2018; Dai et al., 2019, 2017; Tomé et al., 2020); however, the weight after adult emergence parameter was not included in the validation procedures for OECD 239 and therefore there is no standardized approach for generating the weight of emerged adults. Furthermore, it is not currently technically feasible to generate individual bee weights in laboratories where bees complete their development and emerge as adults in a shared enclosed environment, a common approach used by contract laboratories that routinely conduct these studies to support pesticide safety evaluations. Relatively little is known concerning the biological significance and relevance of pollinator protection goals of the weight after adult emergence from in vitro-reared adults. Our second objective was to analyze all available weight data of emerged adults and compare them to the survival endpoints (e.g., NOED and ED50) to determine whether the weight after adult emergence is the most sensitive indicator of a pesticide effect on developing honey bees and therefore assess the value of including weight within the standardized 22-day larval study design. These two research objectives are intended to determine the optimal approach for the generation and interpretation of honey bee larval toxicity data to inform the requirements for a comprehensive yet efficient pesticide risk assessment for pollinators.

METHODS

Study data were acquired from pesticide companies (i.e., registrants) that generated larval toxicity data as part of the data requirements necessary to inform a comprehensive ERA for product registration. All data were anonymized and a study key was developed as a tool to keep the compiled data blind with confidential information (for further information see Supporting Information Appendix A). Numerous study reports from different compounds were compiled, and after triage, a total of 43 compounds were selected based on the availability of endpoints from both study designs (acute and chronic). This selection was made to primarily address the first goal of this project, which involves the comprehensive analysis and comparison of endpoints. Additionally, 46 compounds were selected to meet the requirements of the second goal of this project, as their corresponding study reports from chronic studies included weight data after emergence.

Comparison of endpoints (LD50/LC50) between the acute (seven-day single-dose) and chronic (22-day repeated-dose) larval studies

Anonymized data from acute and chronic honey bee larval studies were requested from pesticide registrants, which included class of pesticide, relevant endpoints, year of the study conduct, and USEPA MRID (Master Record Identifiers) numbers along with additional notes that were pertinent for this evaluation. A total of 43 pesticides (active ingredients or solo formulations), which had both acute and chronic exposure larval toxicity studies (86 studies total), were used in the analysis. All three major classes of pesticides (i.e., herbicides, fungicides, and insecticides) were represented in the data set, including 15 insecticides, 17 fungicides, 10 herbicides, and one herbicide safener (Table 1). Of the available studies, all 43 pesticides had LD50 data from both acute and chronic study designs for comparison, while 41 pesticides (14 insecticides, 16 fungicides, 10 herbicides, and one herbicide safener) were able to provide LC50 data from both acute and chronic study designs for comparison. With the compiled data set, we conducted a comparison of LD_{50} and LC_{50} values from acute study designs following OECD TG 237 and chronic studies following OECD GD 239.

Nondefinitive (greater than) LD50/LC50 endpoints were common for both acute and chronic studies (21 and 25 out of 43 studies, respectively) as a result of either lack of toxicity (LD50 > 100 μg a.i./larva) or difficulty in achieving a

 DOI: 10.1002/ieam.4982

TABLE 1 D7/D8 LD_{50} (μg a.i./larva) and LC_{50} (mg ai/kg-diet) endpoints for both acute and chronic exposure larval toxicity studies

Class	Comp ID	LD50 from acute-D7	LD50 from chronic-D8*	Difference#	LC50 from acute-D7	LC50 from chronic-D8	Difference#
Insecticides	I-1	>15.4	>0.0053	NA	>454	>0.04	NA
	I-2	3.1	>1.2	<2.6	93.94	>7.8	<12
	I-3	0.0128	~0.02	1.6	0.388	~0.127	3.1
	I-4	1.6	> 1.6	> 1.0	52	>11	<4.7
	I-5	0.0012	0.00025	4.8	0.037	0.0015	25
	I-6	1	>1	>1.0	30	>6.33	<4.7
	I-7	0.08	>0.154	>1.9	2.42	>1	<2.4
	I-8	>0.030	0.045	<1.5	>0.909	0.291	>3.1
	I-9	55.9	30	1.9	1650	200	8.3
	I-10	>3.3	>15	NA	>110	>100	NA
	I-11	0.81	0.987	1.2	NR	6.41	NA
	I-12	4.9	>0.4	<12	144	>2.6	<55
	I-13	>30	>24.6	NA	>909	>160	NA
	I-14	0.88	2.205	2.5	26	14.32	1.8
	I-15	0.0539	>0.0183	<2.9	1.6	>0.1188	<14
Fungicides	F-1	23	26	1.1	670	170	3.9
	F-2	>8.07	>3.84	NA	>238	>24.3	NA
	F-3	43.9	>50	>1.1	1295	>324.8	<4.0
	F-4	>30	>50	NA	>914.6	>317	NA
	F-5	>99.2	146	<1.5	>2926	947	>3.1
	F-6	9.8	10.3	1.0	NR	69.4	NA
	F-7	>100	>100	NA	>3030	>650	NA
	F-8	>100	63	>1.6	>3000	380	>7.9
	F-9	67	38	1.8	2000	220	9.1
	F-10	<6.25	>4	>1.6	<189.4	>26	<7.3
	F-11	>80	~40	>2.0	>2300	~250	9.2
	F-12	>50	>25	NA	>1476	>159	NA
	F-13	11.4	>20	>1.8	345.4	>130	<2.7
	F-14	>60	>60.1	NA	>1818	>390	NA
	F-15	>100	>62.5	NA	>3030	>406	NA
	F-16	>100	>33	NA	>3030	>214	NA
	F-17	58.84	>80.1	>1.4	1783	>520	<3.4
Herbicides	H-1	63.36	>27	<2.3	1920	>170	<11
	H-2	45	21	2.1	1300	130	10
	H-3	>100	43	>2.3	>3000	270	>11
	H-4	>5.34	17.6	<3.3	>158	111	>1.4
	H-5	5.8	5.6	1.0	170	35	4.9

(*Continued*)

TABLE 1 (*Continued*)

Class	Comp ID	LD50 from acute-D7	LD50 from chronic-D8*	Difference#	LC50 from acute-D7	LC50 from chronic-D8	Difference#
	H-6	31	48	1.5	900	**290**	**3.1**
	H-7	>10	>23.4	NA	>303	>152	NA
	H-8	>100	>25.1	NA	>3030	>163	NA
	H-9	66.8	> 16	<4.2	2024	>101	<20
	H-10	>100	77	>1.3	>3030	**~500**	**>6.1**
Safener	S-1	>3.3	>7.39	NA	>97	>46.7	NA

Note: Highlighted, in bold (yellow background), the endpoints that are lower (twofold or more), and therefore more toxic, than the endpoint from the comparative study for the designated pesticide.

pesticide concentration in larval diet high enough to elicit toxicity due to low solubility of the pesticide. The LD50/LC50 values were considered different when twofold lower/higher or if the definitive value was lower than the non-definitive value. The twofold difference agrees with USEPA's requirements for reporting adverse effects (i.e., "at levels 50% or more lower than previous acute toxicity studies with similar species") under FIFRA 6(a)(2) (40 CFR § 159.152).

An analysis was conducted to assess the impact of these endpoints on risk assessment using the BeeREX model and USEPA's Tier 1 risk assessment approach (USEPA, 2014) for each pesticide active ingredient in the database. The purpose of this analysis was to determine the extent to which differences in toxicity endpoint values affected the likelihood of a pesticide failing or passing the screening based on RQ values, (RQ) $\left(\frac{\text{Exposure}}{\text{Hazard Endpoint}}\right)$ relative to regulatory levels of concern (LoCs), following the USEPA bee risk characterization approach. Without knowing the identity of the pesticide active ingredient, the exposure portion of the RQ was calculated based on theoretical application rates of 0.1, 0.5, and 1.0 lb a.i./A and multiplying this value by 110 to give the exposure (i.e., pesticide residue in pollen and nectar) estimate in mg a.i./kg (ppm). To convert concentration to a dose based on honey bee consumption, this value is multiplied by 0.1236 µg/bee/day, which is the highest total consumption rate of a honey bee larva used within BeeREX to give the total dose in µg a.i./bee. This exposure value expressed as the dose was divided by the LD50 endpoint (the highest dose level was used for nondefinitive endpoints) to give the RQ, which was then compared to the level of concern (LoC) value of 0.4 for acute risk. To pass the Tier 1 assessment, this calculated RQ should be below LoC.

Weight after adult emergence

Forty-six 22-day chronic larval toxicity studies that included measurements for both survival and weight after adult emergence and statistical analyses were voluntarily provided by pesticide registrants. All study data were anonymized and compiled into a Microsoft Excel-based spreadsheet. The statistical analysis results from these studies were also provided so that the endpoint data could be accurately reflected in the database. For those studies where statistical tests were missing or unreported, individual bee weights from the studies were compiled into a separate spreadsheet. For each study, if these data were normally distributed (Shapiro–Wilk normality test p value > 0.01) and had homogeneous variances (Bartlett test p value > 0.01), a one-sided t-test with Bonferroni adjustment was run in CETIS. If the data were nonnormal or non-homogeneous, a one-sided Wilcoxon rank-sum test with Bonferroni–Holm adjustment was run in R software, to control for family-wise error rate in multiple tests and/or comparisons. These test choices are based on USEPA guidance for continuous data (Newman, 2012; USEPA, 2002). The compiled database included multiple classes of pesticides: herbicides (48% (22/46) of studies; includes plant growth regulators), fungicides (26% (12/46) of studies), and insecticides and/or acaricides (26% (12/46) of studies). An open literature search was also conducted; however, the resulting literature either included substantially different study designs (e.g., duration of study, duration of dosing, or amount of diet fed to larvae, etc.) or did not include the weight after adult emergence. Therefore, the literature assessed was excluded from the database for the purpose of this evaluation.

The dataset for weight after adult emergence utilized different studies than the D7 acute and D22 chronic LD_{50}/LC_{50} data comparison. This dataset included studies where the weight after adult emergence was recorded, which is not currently required according to the current OECD guidance (OECD, 2016), and with a focus on the cumulative NOED and LOED (Lowest Observed Effect Dose) endpoints, which are utilized for estimating chronic risk to bees. For this evaluation, we identified the endpoint that yielded the lowest LOED in each study. The LOED represents the most sensitive endpoint in a study where an effect is observed and therefore was selected as the relevant endpoint for the analyses.

RESULTS

Larval studies LD50/LC50 comparison

The comparative analysis across 43 pesticides (86 studies) revealed that in 88% (38/43) of studies, the pesticide $LD50_{50}$ values derived from the chronic study design were similar

 DOI: 10.1002/ieam.4982

(i.e., within a factor of 2) to the LD50 values derived from the acute study design. Overall, few pesticides had either lower (more sensitive) chronic exposure LD50 values or lower acute exposure LD50 values (1/43 vs. 4/43, respectively) with no discernible differences in measurement endpoints when looked at by pesticide class (Figure 2A). In contrast, when comparing LC50 values from both study designs, it was observed that 66% of the evaluated pesticides (27/41) had LC_{50} values derived from both acute and chronic exposure studies that were similar, and for all of the remaining 14 pesticides (34%), the LC50 values for the chronic study designs were lower than the acute study design. This indicates overall greater sensitivity for the chronic study design in terms of discernible differences in endpoints (Figure 2B).

The impact analysis to assess the risk outcome using the USEPA BeeREX assessment tool revealed that the RQ values calculated with endpoints derived from acute and chronic larval study designs had similar percentages of pesticides exceeding the LoC for all three theoretical application rates proposed (Figure 3A, blue bars), with a slightly higher percent above the LoC for chronic study design with an application rate of 1.0 lb a.i./A.

Additionally, it was observed that when using the USEPA method of converting the cumulative LD50 endpoint, from the 22nd chronic study to a daily dose by dividing by 4 (Farruggia et al., 2022), the estimated RQ values were significantly higher and exceeded the levels of concerns (Figure 3A, purple bar), indicating a higher probability of failing the Tier 1 screening assessment, compared to using the reported acute LD50 values, particularly at the higher application rate of 1.0 lb a.i./A. Finally, when comparing RQs calculated using concentrations rather than converting to dose, the chronic LC50 endpoints (Figure 3A, brown bars) result in a higher percentage of LoC exceedance when compared to the acute LC_{50} versus the estimated daily dose approach (dividing the LD_{50} by 4).

The impact of using these toxicity endpoints is more clearly shown when viewing the RQ results by pesticide class. For insecticides (Figure 3B), most pesticides do not pass the Tier 1 assessment (LoC above 0.4) even at lower rates; however, using the acute LC50 endpoints results in fewer exceedances compared to the other endpoints. In contrast, for fungicides (Figure 3C) and herbicides (Figure 3D), nearly all pesticides exceeded the LoC at application rates above 0.5 lb a.i./A using the estimated daily dose LD50 (dividing cumulative dose by 4). Finally, using LC50 values from the chronic studies provides similar screening-level risk results as using the LD50 values but with the added benefit of not requiring additional conversion of exposure values to the daily dose. This is particularly important when determining risk to other taxa besides honey bees (see discussion).

Weight after adult emergence

Among the studies analyzed for the weight after adult emergence, 22% (10/46) were identified as no-effect studies, where none of the evaluated toxicity endpoints showed a statistically significant reduction compared to the control data. Among the remaining 36 studies, the hazard assessment was driven by one of the four measured endpoints (larval, pupal, or adult survivorship and the weight after adult emergence). In 22% (8/36) of these cases, the most sensitive endpoint was the LOED for the weight after adult emergence (Figure 4). On the other hand, significant effects on survival were observed as the most sensitive endpoint in the majority (78% or 28/36) of the studies. Specifically, among these studies, larval survival was most sensitive (44% or 16/36) of the cases, followed by pupal-stage survival in 28% of cases (10/36) of the cases. Survival to adult emergence was the most sensitive endpoint in only two studies (6%, or 2/36). These results demonstrate that the most sensitive endpoints are typically related to survival effects in the early stages of development (larval and pupal stages) rather than the period between the pupal survival assessment and adult emergence.

The most sensitive endpoint groupings were analyzed further by pesticide class (Figure 5). It is important to mention that there is no strong pattern linking a specific pesticide class to a particular effects endpoint across all three categories. However, it is important to mention that while

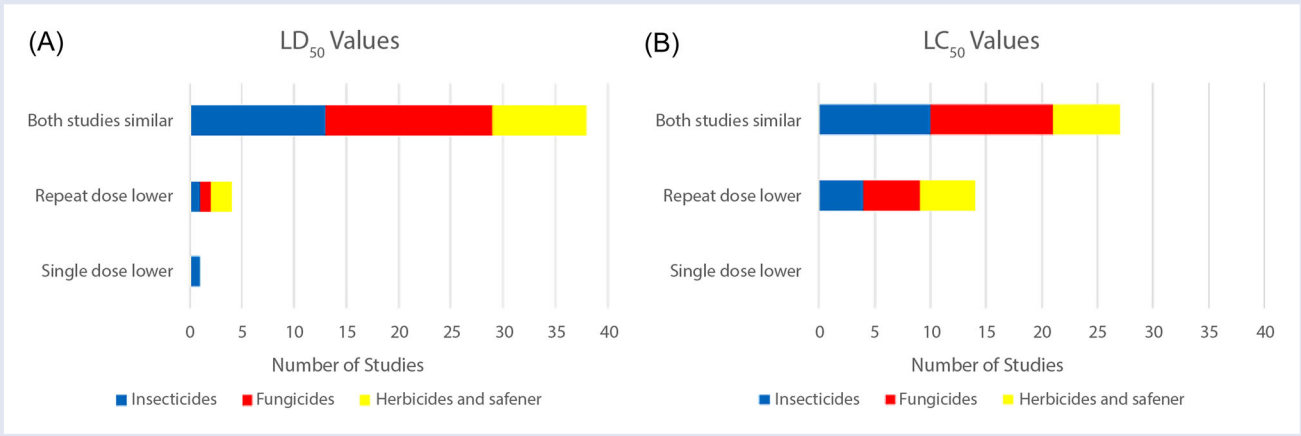

FIGURE 2 Comparison of endpoints (A) LD50 and (B) LC50 derived from chronic study design versus LD/LC50 derived from acute study design at test termination (D8 and D7, respectively)

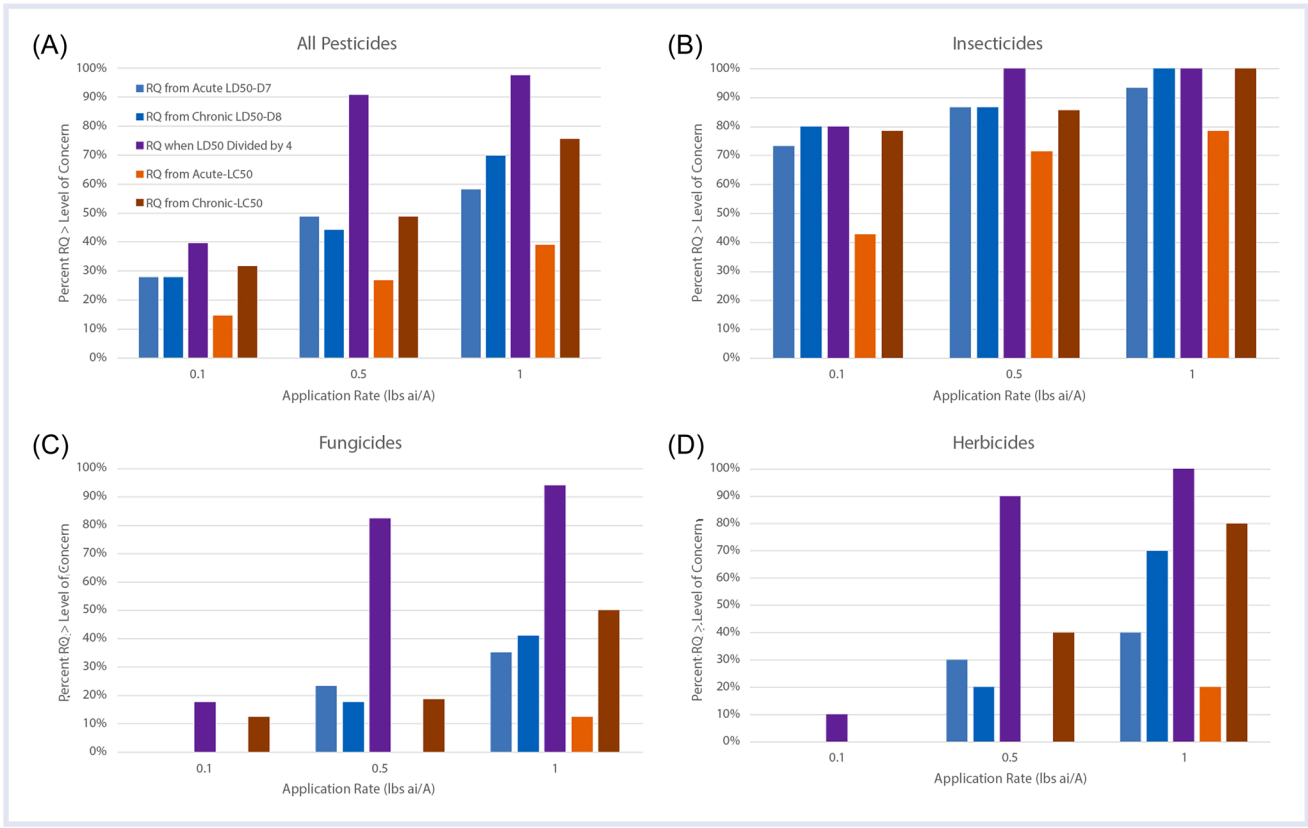

FIGURE 3 Percentage of LD50/LC50 endpoints that would fail USEPA's Tier 1 risk assessment using theoretical applications rates of 0.1, 0.5, and 1.0 lb. a.i./A for (A) all pesticides, (B) fungicides, (C) insecticides, and (D) herbicides

adult survival was not the most sensitive indicator of effects from fungicides, adult weight was impacted in 27% (3/11) of studies where fungicide effects were observed. This suggests that the class of pesticide is not a descriptive factor for predicting the most sensitive study endpoint.

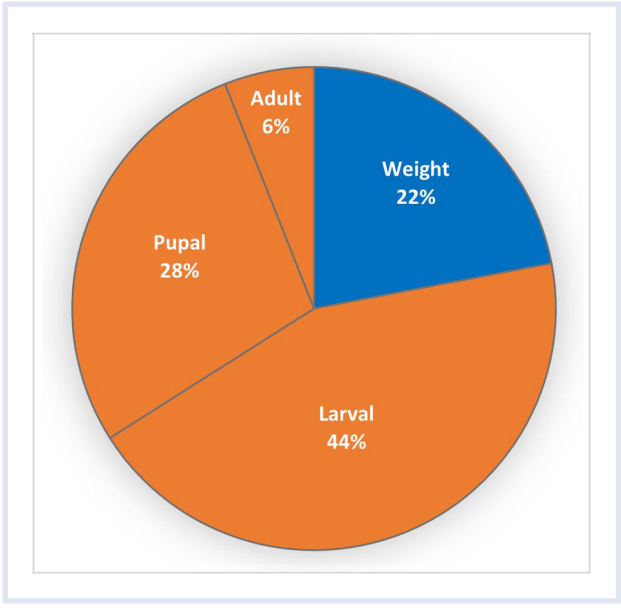

FIGURE 4 Percentage of the 36 studies with an effect, classified by most sensitive/driving endpoint (lowest LOED) in the risk assessment: larval survival, pupal survival, adult emergence, or the weight after adult emergence

The studies that resulted in the weight after adult emergence as the most sensitive endpoint were further analyzed to compare the percentage of mean weight reduction of the significant treatment group (i.e., LOED) to the negative control group. The reduction in body weight after adult emergence ranged from 4% to 12% (Table 2). It is worth noting that in three of these studies (numbers 7, 31, and 36), there were two doses in which weight was more sensitive than the lowest survival or adult emergence LOED and therefore both doses are listed in the table. There was interstudy variability in the patterns of effects observed across the 46 studies in the analysis without any clear trends (see Supporting Information Figure S1) that may indicate difficulty in repeating the weight measurements for a given pesticide across multiple studies (for detailed analysis of biological variability information see Supporting Information Appendix B).

DISCUSSION

Honey bee toxicity study protocols on larvae have now become robust and repeatable tests that enable endpoint assessments at all phases of brood development. Analogous to chronic ecotoxicity studies with other animals, global regulatory agencies rely upon the study data to inform a screening-level assessment of the potential impacts of a pesticide on survival, growth, and reproduction (USEPA, 1998). According to the current bee risk assessment guidance for North American (USEPA, 2014) and Latin American agencies (Cham et al., 2020), both the seven-day acute

 DOI: 10.1002/ieam.4982

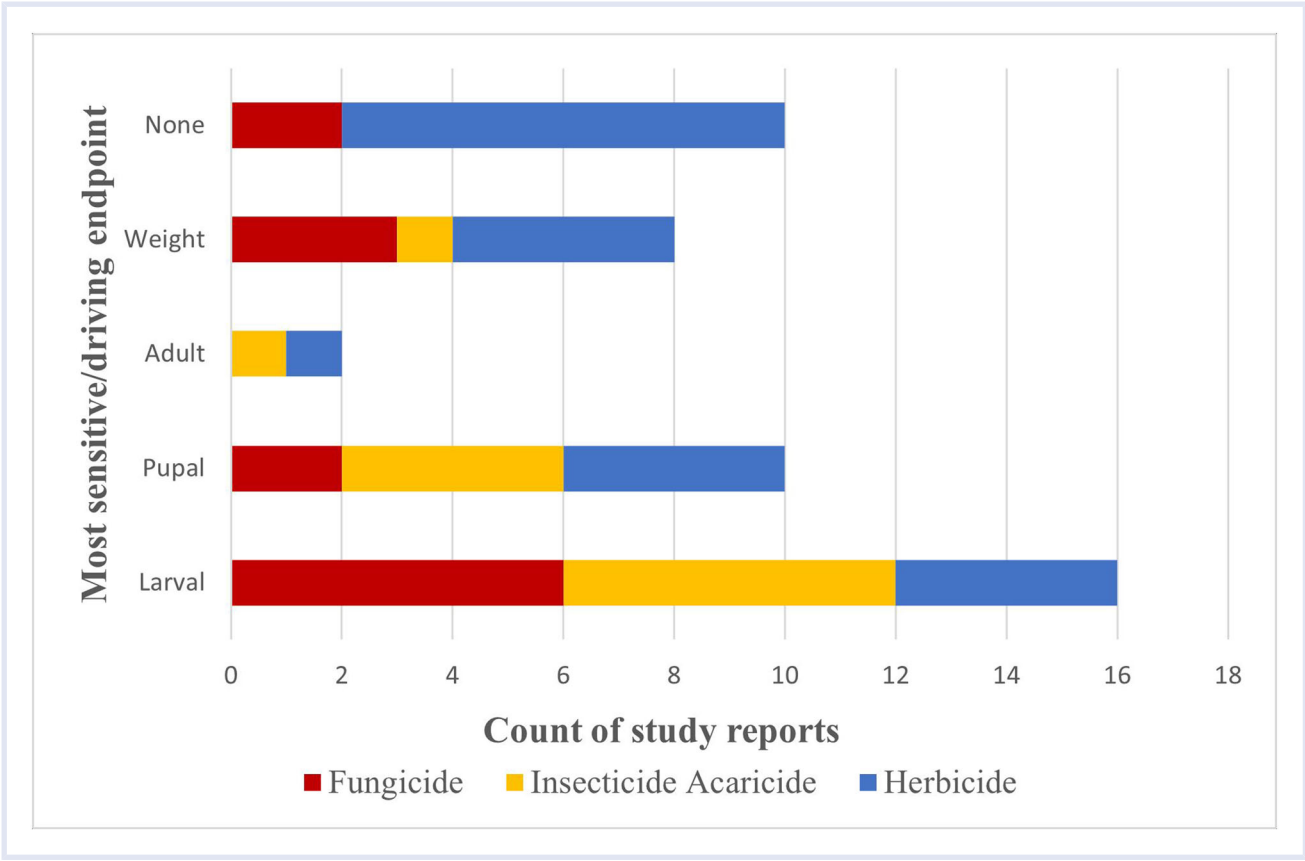

FIGURE 5 Number and pesticide class of studies categorized by most sensitive/driving endpoint

(single-dose) study and a 22-day chronic (repeated-dose) study with honey bee larvae are required for assessing the risk to pollinators at Tier 1 screening phase. The LD50/LC50 survival endpoints from both larval study designs demonstrated that for the majority of studies (88%), LD_{50} endpoints were similar (result less than a two-fold difference). Likewise, the LC50 values are either similar (60%) or lower (34%) (more conservative) in the chronic larval studies when compared to the acute larval study design, indicating that the use of LD50 from acute larval studies is likely less conservative for a screening level assessment and not likely relevant in terms of representing an exposure concentration over a single day

TABLE 2 Mean decreases in weight after emergence in the eight studies in which weight was the most sensitive endpoint

Study	Dose as rank	Negative control group mean weight (g)	Group mean weight of more sensitive doses (g)	Mean weight decrease (g)	Mean weight decrease (% of control)
7	5	0.0943	0.0871	0.0072	7.6
	6	0.0943	0.0833	0.0110	11.7
10	5	0.1041	0.0952	0.0089	8.5
27	3	0.106	0.101	0.005	4.7
28	4	0.1125	0.1009	0.0116	10.3
31	4	0.1128	0.1055	0.0073	6.5
	5	0.1128	0.1019	0.0109	9.7
36	2	0.1060	0.1013	0.0047	4.4
	3	0.1060	0.0998	0.0062	5.8
37	3	0.1041	0.0954	0.0087	8.4
46	5	0.1138	0.1023	0.0115	10.1

during larval development (i.e., diet is not completely consumed over a single day [Aupinel et al., 2009]). Following the Tier 1 assessment scheme through the screening-level tool BeeREX, we evaluated the impact of both LD50 and LC50 endpoints on the RQ based on theoretical application rates. The USEPA typically uses LD50 values in the risk assessment, which requires converting the exposure value (e.g., pollen and nectar residues based on concentration) to a dose (i.e., μg/bee) in calculating the RQ. The benefit of using LC50 values (based on concentration) is that the exposure value does not need to be converted, as both the exposure and effects values are based on concentration, to calculate an RQ. In this analysis, when the LD_{50} values for chronic study design were based on cumulative dose (following OECD 239), we found that the endpoints from both experimental designs yielded similar RQ values. In contrast, when estimating the RQ using the USEPA's method of converting cumulative dose to daily dose (dividing the endpoind value, from cumulative dose by 4, sometimes referred to as LDD50 with units in μg ai/larva/day, per Farruggia et al., 2022), calculated RQ exceeded the level of concern at a much higher rate, indicating that this approach is overly conservative. Furthermore, the approach by Farrugia et al. does not consider the significant variation in food intake that occurs naturally among honey bee larvae during different instar stages (Hartfelder et al., 2015), which ultimately leads to a crude estimate of daily dose and may misinform a screening-level risk assessment for honey bee larvae, triggering unnecessary higher tier refinements.

Acute endpoints based on daily dose are optimal for acute studies; however, it is not technically feasible to determine the precise daily amount of diet consumed by larvae at each developmental stage based on current established protocols of in vitro rearing techniques (Aupinel et al., 2005; Crailsheim et al., 2013; OECD, 2013; Schmehl et al., 2016). The individual larvae consume a total volume of 160 μL diet during the entirety of the OECD 239 study design, yet the actual daily consumption at each larval stage is variable and gradually increases as the developmental stages progress, with the majority of the food intake occurring toward the end of larval development (Haydak, 1970). In the context of risk assessment, the likelihood of exposure, particularly for workers, is highest during the last days of larval development. During the first three days of larval development, worker bees receive a glandular secretion, a milky-white to clear component, known as pure royal jelly (Haydak, 1970; Jung-Hoffmann, 1966). Only from the fourth larval instar onwards, a yellowish component containing pollen grains per se is provided (Haydak, 1970). Within this context, two distinct types of worker jelly can be identified: One is purely secretory in origin and is fed to young larvae, while the other is a mixed diet provisioned to older larvae (Hartfelder et al., 2015). Pesticide contamination levels in royal jelly are known to be low (Milone & Tarpy, 2021), and the amount of pesticides in the larval food diet is significantly correlated with the amount of pollen grains present in the worker jelly. This amount increases with larval age, ranging from 41 to 4654 pollen grains per milligram of worker jelly (Böhme et al., 2019). Additionally, the larva floats on top of the diet during the first few larval instars, representing both a dietary and contact route of exposure that cannot be fully estimated on a daily basis (Tomé et al., 2020). Therefore, from a biological standpoint, the actual relevant exposure to larvae would be best estimated from cumulative exposure through the duration of the larval development stage.

Based on this analysis, it can be concluded that the single acute larval exposure study does not offer additional information to the risk assessment when the 22-day larval study is available. As the current protocol for the 22-day chronic exposure larval study (OECD 239 Guidance Document, OECD, 2016) has been validated through ring-testing, the older single-dose acute larval study has become redundant and unnecessary. The acute larval study design does not accurately provide the daily dose due to the rarity of completely consumed diets within the 24-hour period (as outlined in OECD 237 Guideline), and it does not represent a realistic exposure scenario for larvae within the colony environment. The toxicity values from the chronic larval study tend to be similar or lower (i.e., higher toxicity) than the acute larval study and therefore provide a more robust screening-level risk determination to address the risk posed by a pesticide to the immature stage of bee development (Farruggia et al., 2022; Hilton et al., 2019; Höfer et al., 2004; Wheeler, 2019).

Sublethal effects

The weight after adult honey bee emergence from the 22-day chronic exposure larval assay stands out as a potential indicator for overall growth after spending the developmental stage exposed to a pesticide through both dietary and contact exposures. To address potential effects on growth, the USEPA has requested pesticide registrants to include an assessment of body weight of emerged adult bees as a study endpoint in the 22-day larval chronic study design, despite the absence of a requirement in the OECD guidance document (OECD, 2016) or any technical guidance on how to standardize the collection of body weight from emerged adults. Our study evaluated weight after adult emergence as a study measurement for informing the ERA of a given pesticide. A meta-analysis was conducted using a compilation of assay data from chronic exposure larval toxicity tests, which included measurements of bee body weight after emergence. The analysis identified that survival/adult emergence was the most sensitive endpoint in 78% of the studies that observed some level of effects, whereas the weight after adult emergence was the most sensitive indicator of an effect in the other 22% of the studies.

The results of this assessment showed that survival was the driving indicator of toxic response in most of the studies. However, in the cases where weight was identified as the most sensitive endpoint, the observed mean weight reduction ranged from 4% to 12%. It is unclear from our meta-analysis how the weight measurements were collected, and this may have contributed to the differences observed from our analysis. The determination of the adult's final body weight in

 DOI: 10.1002/ieam.4982

holometabolous insects occurs during the early stages of metamorphosis, which sets the adult size (Colombani et al., 2005; Mirth & Riddiford, 2007). Although significant progress has been made in developing protocols for controlled environment larval rearing, the current protocol methods still provide an approximation of what occurs within a hive, without fully replicating the social context of the in-hive environment (De Souza et al., 2018). Therefore, in vitro-reared bee body weight measurement after emergence can introduce a range of variability due to their innate phenotypic plasticity. It is imperative that methods are standardized for measuring the weight after adult emergence if growth measurements are to be collected, including, but not limited to, the timing of measurements, maintaining the identification of individual bees throughout the entire study duration, fresh versus dry weight, and whether feeding may occur after adult emergence. The continued lack of standardization may have a large impact on the interpretation of the data and its appropriate use within a pesticide risk assessment. It is also critical to understand what level of weight reduction may be relevant when assessing risk to an organism due to normal biological variability within the test system, and whether it changes the conclusions of a risk assessment.

It is important to note that the statistical evaluation producing the lowest LOED value does not always indicate biological relevance, or that the effect was treatment-related. Tests conducted to assess the distribution shape revealed that the weights in the dose groups were not normally distributed and exhibited unequal variances. In such cases, the USEPA (Newman, 2012; USEPA, 2012) recommends the Wilcoxon Rank-Sum test, which indicated that only the highest dose showed significance. However, the OECD (OECD, 2006) recommends the Jonckheere–Terpstra test, which assumes monotonicity and indicates that the three highest doses were significant. For consistency in the current analyses, the originally chosen Jonckheere-Terpstra test was retained, with the significant result of the highest dose added. It should be noted, however, that the choice of test in this specific scenario impacted the resulting LOED. If a different test had been used, the weight after adult emergence would no longer be considered the most sensitive endpoint for this study.

Our analysis demonstrates the suitability of a single 22-day chronic exposure study for assessing the toxicity of a pesticide to developing honey bees. Toxicity values that are based upon a cumulative dose or concentration are the most robust indicator of exposure to inform a comprehensive pesticide risk assessment. The additional benefit of using endpoints based on concentration is that risk determinations can be more easily extrapolated to other immature bee taxa that may differ in the proportion of nectar and pollen consumed during development (Boyle et al., 2018). Additionally, our analysis demonstrated that the weight after adult emergence from the 22-day chronic exposure larval study may be an indicator of pesticide effects on growth and development. However, it needs to be considered with caution as several shortcomings in terms of methodology were identified. Significant discrepancies exist among the documented procedures, ranging from measuring the weight on the day of emergence to trapping and weighing surviving adult bees until Day 22. It is unclear whether these discrepancies in methodology may have impacted our conclusions and emphasize the need to standardize the data collection of the weight after adult emergence within the current study guidance, which is based on a multiple-laboratory test method validation (OECD, 2013, 2016).

AUTHOR CONTRIBUTION

Daiana A. De Souza: Conceptualization; data curation; project administration; visualization; writing—original draft; writing—review and editing. **Max Feken**: Conceptualization; data curation; formal analysis; funding acquisition; methodology; supervision; writing—review and editing. **Hudson V. V. Tomé**: Conceptualization; investigation; writing—review and editing. **Daniel R. Schmehl**: Conceptualization; funding acquisition; methodology; supervision; visualization; writing—review and editing.

ACKNOWLEDGMENT

The data organization and statistical analysis of data presented here was funded by the Pollinator Research Task Force with study data provided by member companies, including BASF Corporation, Bayer CropScience LP, Corteva Agriscience, Mitsui Chemicals Agro, Inc, Syngenta Crop Protection, LLC, UPL NA, Inc., Sumitomo Chemical Company, Ltd., and Valent U.S.A, LLC. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. In addition to the authors, the following individuals contributed to the projects presented here and/or the review of this manuscript: Ana Cabrera, Steve Levine (Bayer CropScience LP), Bridget O'Neil (Corteva Agriscience), Silvia Hinarejos (Sumitomo Chemical Company, Ltd.) and Mark Cafarella (Three Points Ecotox Consulting, LLC). Moreover, we thank Farah Abi-Akar, Jennifer Collins and Jennifer Jackson of Waterborne Environmental, Inc. for their support on drafting the initial manuscript covering the weight after emergence analysis.

CONFLICT OF INTEREST

The authors declare the following financial interests/personal relationships, which may be considered potential competing interests: Daiana A. De Souza reports that financial support, administrative support, and article publishing charges were provided by Pollinator Research Task Force, LLC.

DATA AVAILABILITY STATEMENT

Data, associated metadata, and calculation tools are available from corresponding author Daiana A. De Souza (daiana.almeida-de-souza@basf.com).

ORCID

Daiana A. De Souza http://orcid.org/0000-0003-3499-6478
Daniel R. Schmehl http://orcid.org/0000-0003-1591-8722

SUPPORTING INFORMATION

Additional description of methods and supplemental results.

REFERENCES

Aupinel, P., Fortini, D., Dufour, H., Tasei, J.-N., Michaud, B., Odoux, J.-F., & Pham-Delègue, M.-H. (2005). Improvement of articial feeding in a standard in vitro method for rearing *Apis mellifera* larvae. *Bulletin of Insectology*, *58*(2), 107–111.

Aupinel, P., Fortini, D., Michaud, B., Mędrzycki, P., Padovani, E., Przygoda, D., Maus, C., Charrière, J.-D., Kilchenmann, V., Riessberger-Gallé, U., Vollmann, J., Jeker, L. T., Janke, M., Odoux, J.-F., & Taséi, J. N. (2009). Honey bee brood ring-test: Method for testing pesticide toxicity on honeybee brood in laboratory conditions. *Julius-Kühn-Archiv*, *423*, 96–102.

Böhme, F., Bischoff, G., Zebitz, C. P. W., Rosenkranz, P., & Wallner, K. (2019). From field to food II—Will pesticide-contaminated pollen diet lead to a contamination of worker jelly? *Journal of Apicultural Research*, *58*, 542–549.

Boyle, N. K., Pitts-Singer, T. L., Abbott, J., Alix, A., Cox-Foster, D. L., Hinarejos, S., Lehmann, D. M., Morandin, L., O'Neill, B., Raine, N. E., Singh, R., Thompson, H. M., Williams, N. M., & Steeger, T. (2018). Workshop on pesticide exposure assessment paradigm for non-Apis bees: Foundation and summaries. *Environmental Entomology*, *48*, 4–11.

Bruce, A. B., Borlu, Y., & Glenna, L. L. (2023). Assessing the scientific support for U.S. EPA pesticide regulatory policy governing active and inert ingredients. *Journal of Environmental Studies and Sciences*, *13*, 1–13.

Cham, K. de O., Rebelo, R. M., Oliveira, R. de P., Ferro, A. A., Viana Silva, F. E. de C., Borges, L. de O., Saretto, C. O. S. D., Tonelli, C. A. M., & Macedo, T. C. (2020). *Manual de avaliação de risco ambiental de agrotóxicos para abelhas* (2nd ed.). Ibama/Diqua.

Clark, S. L. (2017). *Results of the honey bee* (Apis mellifera *L.*) *toxicity following chronic (repeated) exposure international ring test performance*. (Technical Report Prepared for the Pollinator Research Task Force). https://pollinatorresearchtaskforce.com/wp-content/uploads/2019/03/PRTF-Projects-Ring-Study-Report-Feb-2017.pdf

Colombani, J., Bianchini, L., Layalle, S., Pondeville, E., Dauphin-Villemant, C., Antoniewski, C., Carré, C., Noselli, S., & Léopold, P. (2005). Antagonistic actions of ecdysone and insulins determine final size in Drosophila. *Science*, *310*, 667–670.

Crailsheim, K., Brodschneider, R., Aupinel, P., Behrens, D., Genersch, E., Vollmann, J., & Riessberger-Gallé, U. (2013). Standard methods for artificial rearing of *Apis mellifera* larvae. *Journal of Apicultural Research*, *52*, 1–16.

Dai, P., Jack, C. J., Mortensen, A. N., Bloomquist, J. R., & Ellis, J. D. (2018). The impacts of chlorothalonil and diflubenzuron on *Apis mellifera* L. larvae reared in vitro. *Ecotoxicology and Environmental Safety*, *164*, 283–288.

Dai, P., Jack, C. J., Mortensen, A. N., Bustamante, T. A., Bloomquist, J. R., & Ellis, J. D. (2019). Chronic toxicity of clothianidin, imidacloprid, chlorpyrifos, and dimethoate to *Apis mellifera* L. larvae reared in vitro. *Pest Management Science*, *75*, 29–36.

Dai, P., Jack, C. J., Mortensen, A. N., Bustamante, T. A., & Ellis, J. D. (2018). Chronic toxicity of amitraz, coumaphos and fluvalinate to *Apis mellifera* L. larvae reared in vitro. *Scientific Reports*, *8*, 5635.

Dai, P., Jack, C. J., Mortensen, A. N., & Ellis, J. D. (2017). Acute toxicity of five pesticides to *Apis mellifera* larvae reared in vitro. *Pest Management Science*, *73*(11), 2282–2286.

De Souza, D. A., Kaftanoglu, O., De Jong, D., Page, Jr., R. E., Amdam, G. V., & Wang, Y. (2018). Differences in the morphology, physiology and gene expression of honey bee queens and workers reared in vitro versus in situ. *Biology Open*, *7*(11), bio036616.

Farruggia, F. T., Garber, K., Hartless, C., Jones, K., Kyle, L., Mastrota, N., Milone, J. P., Sankula, S., Sappington, K., Stebbins, K., Steeger, T., Summers, H., Thompson, P. G., & Wagman, M. (2022). A retrospective analysis of honey bee (*Apis mellifera*) pesticide toxicity data. *PLoS One*, *17*, e0265962.

Hartfelder, K., Guidugli-Lazzarini, K. R., Cervoni, M. S., Santos, D. E., & Humann, F. C. (2015). Chapter One—Old threads make new tapestry—Rewiring of signalling pathways underlies caste phenotypic plasticity in the honey bee, *Apis mellifera* L. In A. Zayed & C. F. Kent (Eds.), *Advances in insect physiology* (Vol. 48, pp. 1–36). Academic Press.

Haydak, M. H. (1970). Honey bee nutrition. *Annual Review of Entomology*, *15*, 143–156.

Hilton, G. M., Odenkirchen, E., Panger, M., Waleko, G., Lowit, A., & Clippinger, A. J. (2019). Evaluation of the avian acute oral and sub-acute dietary toxicity test for pesticide registration. *Regulatory Toxicology and Pharmacology*, *105*, 30–35.

Höfer, T., Gerner, I., Gundert-Remy, U., Liebsch, M., Schulte, A., Spielmann, H., Vogel, R., & Wettig, K. (2004). Animal testing and alternative approaches for the human health risk assessment under the proposed new European chemicals regulation. *Archives of Toxicology*, *78*, 549–564.

Johnson, R. (2014). Honey bee toxicology. *Annual Review of Entomology*, *7*(60), 415–434.

Jung-Hoffmann, L. (1966). Die Determination von Königin und Arbeiterin der Honigbiene. *Zeitschrift fuer Bienenforschung*, *8*, 296–322.

Maus, C., Alix, A., & Schmehl, D. (2022). Assessing the impact of pesticides on pollinators. In *Promoting pollination and pollinators in farming* (pp. 183–224). https://doi.org/10.19103/AS.2022.0111.14

Milone, J. P., Chakrabarti, P., Sagili, R. R., & Tarpy, D. R. (2021). Colony-level pesticide exposure affects honey bee (*Apis mellifera* L.) royal jelly production and nutritional composition. *Chemosphere*, *263*, 128183.

Milone, J. P., & Tarpy, D. R. (2021). Effects of developmental exposure to pesticides in wax and pollen on honey bee (*Apis mellifera*) queen reproductive phenotypes. *Scientific Reports*, *11*(1), 1020.

Mirth, C. K., & Riddiford, L. M. (2007). Size assessment and growth control: How adult size is determined in insects. *BioEssays*, *29*, 344–355.

Newman, M. C. (2012). *Quantitative ecotoxicology*. CRC Press.

OECD. (2006). *Current approaches in the statistical analysis of ecotoxicity data*.

OECD. (2013). *Test No. 237: Honey bee (Apis Mellifera) larval toxicity test, single exposure*.

OECD. (2016). *Guidance document No. 239: Honey bee larval toxicity test following repeated exposure* (ENV/JM/MONO).

Sanchez-Bayo, F., & Goka, K. (2014). Pesticide residues and bees—A risk assessment. *PLoS One*, *9*, e94482.

Schmehl, D., Tomé, H. V. V., Mortensen, A., Martins, G., & Ellis, J. (2016). Protocol for the in vitro rearing of honey bee (*Apis mellifera* L.) workers. *Journal of Apicultural Research*, *55*(2), 113–129.

Siviter, H., Linguadoca, A., Ippolito, A., & Muth, F. (2023). Pesticide licensing in the EU and protecting pollinators. *Current Biology*, *33*, R44–R48.

Tomé, H. V. V., Schmehl, D., Wedde, A., Maia, R., Ravaiano, S., Guedes, R. N., Martins, G., & Ellis, J. (2020). Frequently encountered pesticides can cause multiple disorders in developing worker honey bees. *Environmental Pollution*, *256*, 113420.

USEPA. (1998). *Guidelines for ecological risk assessment*.

USEPA. (2002). *Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms* (EPA 821-R-022-013).

USEPA. (2012). *White paper in support of the proposed risk assessment process for bees*. Submitted to the FIFRA Scientific Advisory Panel for Review and Comment September 11–14, 2012. EPA Office of Chemical Safety and Pollution Prevention Office of Pesticide Programs Environmental Fate and Effects Division, Environmental Assessment Directorate, Pest Management Regulatory Agency, Health Canada, California Department of Pesticide Regulation. https://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2012-0543-0004

USEPA. (2014). *Guidance for assessing pesticide risks to bees*. Office of Pesticide Programs, US Environmental Protection Agency. https://www.epa.gov/sites/production/files/2014-06/documents/pollinator_risk_assessment_guidance_06_19_14.pdf

USEPA. (2016). *Guidance on exposure and effects testing for assessing risks to bees*. Office of Pesticide Programs, US Environmental Protection Agency. https://www.epa.gov/sites/production/files/2016-07/documents/guidance-exposure-effects-testing-assessing-risks-bees.pdf

Wheeler. (2019). *Memorandum: Directive to prioritize efforts to reduce animal testing*.

