JOURNAL OF TOXICOLOGY AND ENVIRONMENTAL HEALTH, PART B
2017, VOL. 20, NOS. 6–7, 383–386
https://doi.org/10.1080/10937404.2017.1388574

Taylor & Francis
Taylor & Francis Group

Quantitative weight of evidence assessment of risk to honeybee colonies from use of imidacloprid, clothianidin, and thiamethoxam as seed treatments: a postscript

Keith R Solomon[a] and Gladys L Stephenson[b]

[a]Centre for Toxicology, School of Environmental Sciences, University of Guelph, Guelph, Canada; [b]Aquaterra Environmental Consulting Inc., Orton, Canada

ABSTRACT

This paper is a postscript to the four companion papers in this issue of the Journal (Solomon and Stephenson 2017a, 2017b; Stephenson and Solomon 2017a, 2017b). The first paper in the series described the conceptual model and the methods of the QWoE process. The other three papers described the application of the QWoE process to studies on imidacloprid (IMI), clothianidin (CTD), and thiamethoxam (TMX). This postscript was written to summarize the utility of the methods used in the quantitative weight of evidence (QWoE), the overall relevance of the results, and the environmental implications of the findings. Hopefully, this will be helpful to others who wish to conduct QWoEs and use these methods in assessment of risks.

The use of data from higher tier studies

The entire focus of this QWoE was to assess the evidence for effects of neonicotinoids on honeybees at the level of the colony. In nature, the honeybee does not exist as a solitary organism. Like many colonial insects (bees, ants, wasps), the honeybee is a colony of different castes of the same organism, all working together to ensure the sustainability of the colony and its members. The castes in a honeybee colony are the queen, the drones, and the worker bees (nurse bees, water carriers, and forager bees). There is redundancy and resilience in the castes except for the queen. The overall performance of the colony has been recommended as an indicator of sustainability for the purposes of risk assessment (USEPA 2014). Therefore, we assessed primarily studies that included whole-colony toxicity tests with food sources (sugar solution and/or pollen) containing a range of concentrations of the insecticide as the sole source of food. These studies provided data on several colony-level endpoints and were used to identify the no observed adverse effect concentration (NOAEC) and lowest observed adverse effect concentration (LOAEC). For comparison to other routes of exposure, these were converted to dose per bee. These data were then compared to field studies conducted under similar conditions to the toxicity tests where the bees were confined to food sources from plants grown from seeds treated with neonicotinoids and effects characterized or concentrations in the food- and water-sources for bees assessed. Thus, laboratory-toxicity data for individual bees were not used in the process and extrapolations from individual bees to effects at the level of the colony were not needed.

Selection of papers and reports

It is important to recognize that all available papers and reports were used in the QWoEs. Quality of studies has frequently been used to select the best toxicity data for either assessment of risk or setting of criteria. Guidelines for assessment of quality of data have been suggested, (e.g., see Breton et al. 2009; Hanson et al. 2017; Klimisch, Andreae, and Tillmann 1997; Moermond et al. 2017, 2016; USEPA 2011) but in almost all of these, the assessment is used to select the best-quality data and exclude the lesser-quality data. The selection of papers and

CONTACT Keith R Solomon ksolomon@uoguelph.ca Centre for Toxicology, School of Environmental Science, University of Guelph, Guelph, ON, Canada N1G 2W1.

reports that we used in QWoE was not exclusionary. All papers or reports were included in the WoE process unless there were insufficient data in the report or paper to allow quality to be assessed. The philosophy behind this decision was that some preliminary studies of lesser quality might contain observations that illuminated new toxic mechanisms and/or adverse effects not considered previously. If identified in this way, these studies could be used to evaluate uncertainty and point to alternative interpretations that could warrant additional investigation.

Developing the scoring guides

When assessing quality and relevance of data, many factors must be considered (Hanson et al. 2017; Moermond et al. 2017). These factors are used to establish criteria for evaluation of studies in the QWoE. In most cases, these criteria can be developed from study-guidelines such as those published by the Organization of Economic Cooperation and Development (OECD 2017). However, because there were no OECD guidelines for studying the effects of chemicals on honeybee colonies, we had to develop our own criteria. The basis for our criteria was prior experience in similar QWoEs (Bridges and Solomon 2016; Van Der Kraak et al. 2014), the application of good science, appropriate experimental designs, and the reading of several papers and reports on studies on honeybee-colonies. Scoring guidelines were finalized *a priori* as any changes in the scoring during the process of evaluation (which in our example took about six months for each chemical) must be reapplied to all completed studies. This presented a challenge; it was impossible to anticipate all strengths and weakness in reports that had not been published. Because of this, we used a score for expert judgment (SEJ), which allowed us to incorporate scores for unanticipated weaknesses or strong points in the design of the study.

Enumeration of the quality of the methods used

General and critical criteria (see discussion in Moermond et al. 2017) were used to enumerate quality of the methods (QoM) used in the studies. The general criteria were designed to capture generic components of a good experimental design and, when these were missing, they were identified as weaknesses and resulted in a reduction in the maximum score. Where major weaknesses were identified, these resulted in a greater reduction (25%) in the score. Critical criteria were also used in the assessment of quality and were more specific to the objectives of the types of studies. Scores were based on specific items in the methods such as the number of replicates, the number of sites, the sampling period. Others related to the reporting of the study and included the clarity and details of the methods, confirmation of exposures, and the transparency of the data. Reports conducted under Good Laboratory Practice (GLP) guidelines generally scored better in the description of the methods and the transparency of the data. Published papers scored lower because methods were often incompletely reported and the raw data were generally not provided. Every weakness and SEJ score was documented in the description of the study and this was included in the supplemental information for the papers on IMI, CTD, and TMX. The WoEs underwent a quality assurance check before scores were processed into the graphics. This step was very useful as it identified data errors and inconsistencies in scoring of different studies. Although general criteria for scoring were more subjective than critical criteria, the scoring guides were the same for all evaluations and all the studies were assessed using the same criteria with equal rigor.

Scaling of quality and relevance of the scores

The scale of the scores for quality and relevance was from 0 to 4. As was done in a previous QWoE (Van Der Kraak et al. 2014), scores for each criterion were restricted to whole-numbers. However, the overall score for the QoM was an arithmetic mean to one significant figure. The scale (0–4) was arbitrary and the sole function of this was to separate the higher- from the lower-quality studies. These values have no relationship to scores used in data selection schemes (Breton et al. 2009; Klimisch, Andreae, and Tillmann 1997; Moermond et al. 2017, 2016); however, they could be used for that purpose if appropriate cut-

off values were selected. Likewise, the score for relevance was from 0–4 and was only used to separate observations that were relevant to adverse effects from those that were not. The scores for relevance were based on statistical significance or, for concentrations measured in bee-relevant matrices, by comparison to toxicity values derived from tests conducted at the level of the colony.

Causality

The QWoE process did not allow for the assignment of causality in the sense of most of the Bradford-Hill guidelines, (Hill 1965) but, for experimental exposures in tunnels (tents) where bees were forced to feed on nectar and/or pollen from treated plants, the lack of effects enumerated in the QWoEs provides evidence that is applicable to the guideline of consistency. Most of the studies on the potential effects of neonicotinoids at the level of the hive were semi-field (tunnel) studies where crops were treated with neonicotinoids as well as fungicides at the recommended label rates so the guideline of biological gradient could not be tested. However, one study on TMX (Syngenta 2001), described in more detail in the supplemental information of (Stephenson and Solomon 2017b) employed rates of seed treatment of 1, 2, 3, 4, 6, and 8-times the normal field rate (420 g/100 kg seed) for spring oil-seed rape. The study had several weaknesses (QoM = 0.8) the major one being a lack of replication, but the regression design, the equivalent of a concentration-response experiment, did address the Bradford-Hill guideline of biological gradient. Of the six colony-level endpoints measured in this study, only flight intensity in honeybees confined to the treated crop at 8-times the rate of application was reduced, indicating a margin of exposure of 6-fold. A more robust design based on this regression design could provide better information for assessing risks.

Hazard, risk, and the probability of alternative conclusions

Hazard is the property of a chemical that can result in harm. For the neonicotinoids, hazard to pest organisms is desired, but hazard to non-targets is not. Whether the pesticide causes harm is dependent on exposure and risk is the probability that an organism will receive a harmful exposure. When characterizing concentrations of neonicotinoids in bee-relevant matrices, our QWoE did consider risk, but not in the fully probabilistic sense of probability of harm (such as in Giddings et al. 2014, and numerous other papers). The toxicity values used in characterising concentrations of neonicotinoids were derived from the NOAEC and LOAEC measured at the level of the colony in experimental studies where honeybees were fed with food amended with insecticide. Use of these values did not incorporate a concentration-response of the colony-level effects, but it did provide bounds of exposure that were used in a deterministic worst-case characterization of risk. This approach is like that widely used in assessment of risks in humans, such as in the calculation of acceptable daily intake or reference dose. However, as honeybee colonies were the object of the assessment as well as the "test organisms", interspecies extrapolations were not needed. Concentrations of neonicotinoids in bee-relevant matrices that were less than the NOAEC were deemed to present no risk (score for relevance to adverse effect = 0), those between the NOAEC and LOAEC presented some risk (score for relevance = 2) and those above the LOAEC were deemed to present risk of adverse effects (score for relevance = 4). For small data sets or where centiles were not or could not be calculated, maximum concentrations of neonicotinoid were used for assessment of risk in a simple deterministic risk quotient (exposure concentration/toxicity value). Where sufficient data were provided, the 90th centile concentration was compared to the NOAEC and LOAEC in a semi-probabilistic determination of risk. As foraging honeybees do not directly consume the nectar and pollen that they collect, they (and the colony) are exposed to the average or the median concentration. Thus, the use of the 90th centile of exposure-concentrations is a reasonable worst case and provides a margin of safety. In field studies where effects of neonicotinoids in honeybee colonies confined to treated crops were evaluated, the consistent lack of significant adverse effects provided a measure of certainty that the likelihood of alternative conclusions (i.e., adverse effects) was small.

Overall, the conclusions for the QWoE assessments of IMI, CTD, and TMX were similar and indicated that use of these insecticides as seed treatments under good agricultural practices (GAP) did not result in harm to honeybees at the level of the colony. Given the slight differences in the colony-level toxicity values and the similarity of the rates of treatment used under GAP, there is an overall consistency in the results. None of these neonicotinoids appear to present more risk that the others. This was true for exposures to measured concentrations in bee-relevant matrices as well as in experimental studies where bees were confined to treated crops, both representing a reasonable worst case. Consistency in the observations within and between these three neonicotinoids further supports the general overall conclusion of lack of harm to colonies of honeybees.

References

[OECD] OECD. 2017. OECD guidelines for the testing of chemicals, sections 1-5. Organization for Economic Cooperation and Development Accessed January 2017. http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals_chem_guide_pkg-en

Breton, R. L., G. Gilron, R. Thompson, S. Rodney, and S. Teed. 2009. A new quality assurance system for the evaluation of ecotoxicity studies submitted under the new substances notification regulations in Canada. *Integrated Environmental Assessment and Management* 5:127–37. doi:10.1897/IEAM_2008-026.1.

Bridges, J., and K. R. Solomon. 2016. Quantitative weight of evidence analysis of the persistence, bioaccumulation, toxicity and potential for long range transport of the cyclic volatile methyl siloxanes. *Journal of Toxicogy and Environment Health B* 19:345–79. doi:10.1080/10937404.2016.1200505.

Giddings, J., M. Williams, K. R. Solomon, and J. P. Giesy. 2014. Risks to aquatic organisms from the use of chlorpyrifos in the United States. *Reviews of Environmental Contamination and Toxicology* 231:119–62.

Hanson, M. L., B. A. Wolff, J. W. Green, M. Kivi, G. H. Panter, M. S. J. Warne, M. Ågerstrand, and J. P. Sumpter. 2017. How we can make ecotoxicology more valuable to environmental protection. *The Science of the Total Environment* 578:228–35. doi:10.1016/j.scitotenv.2016.07.160.

Hill, A.B. 1965. The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine* 58:295–300.

Klimisch, H.-J., M. Andreae, and U. Tillmann. 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regulatory Toxicology and Pharmacology* 25:1–5. doi:10.1006/rtph.1996.1076.

Moermond, C., A. Beasley, R. Breton, M. Junghans, R. Laskowski, K. R. Solomon, and H. Zahner. 2017. Assessing the reliability of ecotoxicological studies: An overview of current needs and approaches. *Integrated Environmental Assessment and Management* 13:640–51. doi:10.1002/ieam.1870.

Moermond, C., R. Kase, M. Korkaric, and M. Agerstrand. 2016. CRED: Criteria for reporting and evaluating ecotoxicity data. *Environmental Toxicology and Chemistry* 35:1297–309. doi:10.1002/etc.3259.

Syngenta. 2001. Semi-Field Test: Side Effects of Oil-Seed Spring-Rape (*Brassica napus*) Dressed with Different Rates of CGA 293343 on the Honey Bee (*Apis mellifera* L.). Basel, Switzerland: Syngenta Ltd. (Unpublished Report). Report CGA293343_1833 - A9807C, A9567C (20001077/01-BZEU).

Solomon, K. R., and G. L. Stephenson. 2017a. Quantitative weight of evidence assessment of higher tier studies on the toxicity and risks of neonicotinoid insecticides in honeybees 1: Methods. *Journal of Toxicology and Environmental B* This Issue.

Solomon, K. R., and G. L. Stephenson. 2017b. Quantitative weight of evidence assessment of higher tier studies on the toxicity and risks of neonicotinoid insecticides in honeybees 3: Clothianidin. *Journal of Toxicology and Environmental Health B* This issue.

Stephenson, G. L., and K. R. Solomon. 2017a. Quantitative weight of evidence assessment of higher tier studies on the toxicity and risks of neonicotinoid insecticides in honeybees 2: Imidacloprid. *Journal of Toxicology and Environmental Health B* This issue.

Stephenson, G. L., and K. R. Solomon. 2017b. Quantitative weight of evidence assessment of higher tier studies on the toxicity and risks of neonicotinoid insecticides in honeybees 4: Thiamethoxam. *Journal of Toxicology and Environmental Health B* This issue.

USEPA. 2011. Evaluation guidelines for ecological toxicity data in the open literature. Procedures for Screening, Viewing, and Using Published Open Literature Toxicity Data in Ecological Risk Assessments. Washington, DC: U.S. Environmental Protection Agency, Office of Pesticide Programs. https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/evaluation-guidelines-ecological-toxicity-data-open

USEPA. 2014. Guidance for Assessing Pesticide Risks to Bees. Washington, DC: Office of Chemical Safety and Pollution Prevention, Office of Pesticide Programs. United States Environmental Protection Agency.https://www.epa.gov/sites/production/files/2014-06/documents/pollinator_risk_assessment_guidance_06_19_14.pdf

van der Kraak, G. J., A. J. Hosmer, M. L. Hanson, W. Kloas, and K. R. Solomon. 2014. Effects of atrazine in fish, amphibians, and reptiles: An analysis based on quantitative weight of evidence. *Critical Reviews in Toxicology* 44 (S5):1–66. doi:10.3109/10408444.2014.967836.