Section 1 – Risk Assessment

1.1 Estimating honeybee forager background mortality: a case study in the Netherlands

Ivo Roessink[1], Sjef van der Steen, Dick Belgers, Hans Baveco, Andreas Focks, Jos Boesten

[1] Wageningen Environmental Research (Alterra), P.O. Box 47 6700AA Wageningen, The Netherlands

DOI 10.5073/jka.2018.462.002

Abstract

One of the key assumptions in the EFSA guidance on the risk assessment of plant protection products on bees (2013) concerns the value of honeybee forager background mortality. This background mortality is crucial because its value feeds directly into the trigger value used in the Tier-1 risk assessment. Low forager background mortality results in conservative trigger values, whereas higher forager mortality values result in less conservative triggers. A proper estimate of forager background mortality is therefore key to a realistic and robust risk assessment.

Data underlying the current estimate of forager mortality mostly originate from studies performed outside of Europe, with only one European study being available in the city centre of Basel. The value used in Tier 1 (5.3% mortality per day) is the measurement from Basel because this was the lowest value found. Since the city centre of Basel is not representative for European agricultural environments, a new study was performed that was focussed on the estimation of forager mortality in a realistic agricultural setting in the Netherlands. Freshly emerged honeybees (age <24h) from two hives were tagged every two weeks with micro-transponder RFID chips at the outdoor experimental station 'De Sinderhoeve' . Tagging continued from June to October and every tagged cohort was followed in time. Bees were detected: a) upon tagging, b) when they left the hive and c) when they entered the hive. First results of data evaluation indicate that already within 1 week some bees left the hive briefly but that foraging commenced usually after two weeks and lasted in individual cases for more than 5 weeks after tagging. Based on the obtained data sets, first estimates reveal a honeybee forager background mortality of at least 10% indicating that the EFSA assumption is conservative.

1.2 Three cardinal numbers to safeguard bees against pesticide exposure: LD_{50}, NOEC (revised) and the Haber exponent.

James E. Cresswell

Biosciences, University of Exeter, Exeter EX4 4PS, UK.

DOI 10.5073/jka.2018.462.003

Abstract

Regulators often employ cardinal indicators to justify measures to protect the health of farmland bees from pesticides used in crop protection. Previously, in evaluating the likely hazard of a compound, they have made extensive use of its LD_{50} ('lethal dose to 50% of exposed subjects'), and NOEC ('no observable effect concentration'). Here, I argue that regulators should also use a third indicator, namely the Haber exponent. The Haber exponent qualifies the meaning of the LD_{50} by revealing the relative hazard of environmentally relevant exposures longer than that used to determine the LD_{50} originally. Additionally, I show how the experimental protocol used to determine the Haber exponent will also produce a well-founded, parametric value of the NOEC. Taken together, these three numbers establish a strong foundation on which to evaluate the potential impact of an agrochemical on bees.

Introduction

Regulators need scientific evidence to justify measures to protect the health of farmland bees from pesticides used in crop protection. The best evidence is provided by experiments that closely simulate realistic scenarios, such as field trials that reveal the degree of harm that a pesticide causes to bees when used in farming practice. However, regulators also can make use of cardinal indicators, by which I mean certain numbers whose values carry information about either the comparative toxicity or absolute hazard of an active substance. Two of the cardinal values are

the LD_{50} ('lethal dose to 50% of exposed subjects'), and the NOEC ('no observable effect concentration'). Here, I propose a third: the Haber exponent. Below, I argue that establishing these three numbers for an agrochemical makes a strong foundation on which to evaluate the potential impact of a compound on bees.

First cardinal number: LD_{50}

The LD_{50} is not useful to regulators as a 'protection threshold', or a maximum permissible level of exposure because it safeguards only half of the population, which is not normally sufficient. Useful protection thresholds can, however, be derived from the LD_{50}. For example, regulators may consider imposing a threshold of $LD_{50/10}$, which has the following theoretical justification. The LD_{50} is a percentile on the cumulative distribution of the frequency distribution of dose tolerances in the exposed population. If the frequency distribution of tolerances (minimum lethal doses of toxicant) in the population is normally distributed, then the cumulative distribution is sigmoidal and the LD_{50} coincides with the mode (and mean) of the frequency distribution (Fig. 1). If the population varies little in tolerance, the sigmoidal cumulative distribution rises steeply, otherwise it is shallow. If we require that the frequency distribution of tolerances is unimodal, then a theoretical asymptote arises when the frequency distribution of tolerances is flat (Fig. 1, asymptotic case). Arguably, this is a 'worst case scenario' because some members of the population have no tolerance to the exposure and others have virtually none. In this hypothetical worst case, the cumulative distribution is a straight line (Fig. 1) and it is possible to be very precise about the death rate at an exposure of $LD_{50/10}$; specifically, it is (50/10)% = 5%. Since this is a worst case, we can say that a protection threshold of $LD_{50/10}$ will result in the death of no more than 5% of the population – and less if the distribution of tolerances is unimodal and normal, which is more likely in realistic scenarios. Note that $LD_{50/10}$ becomes more effective as a protection threshold as populations vary less in tolerance.

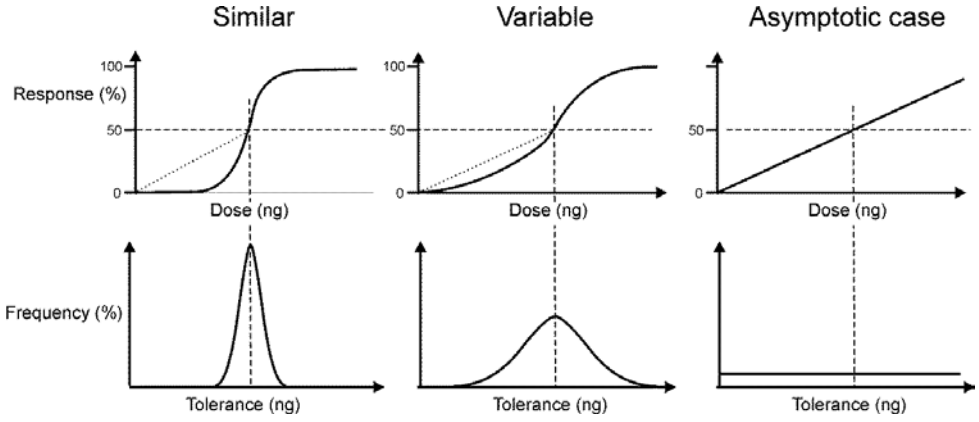

Fig. 1 Relationships between the dose-response curve (upper panels) and the frequency distribution of dose tolerances (lower panels) in each of three populations (the upper and lower panels are paired within the three columns: similar, variable and asymptotic case). The within-population variability among individuals in dose tolerance increases progressively left-to-right across the three columns. Vertical dashed lines indicate the modes of the frequency distributions and the horizontal dashed lines indicate the LD_{50} (i.e. the response endpoint is fatality).

The second use of the LD_{50} is as a comparative indicator. Over time, toxicologists have determined the LD_{50} of many compounds so that by comparing the LD_{50} of a new compound to the known impacts of compounds in past use, regulators can quickly form an opinion about the likely relative hazard of the new one.

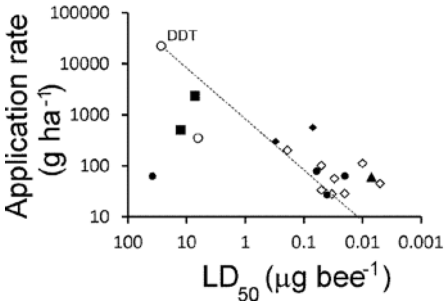

Fig. 2 Relationship between (x-axis) and typical agricultural application rate (y-axis) for a collection of neurotoxic insecticides. ○ denotes DDT; ■ = carbamate, ◇ = pyrethroids, ● = neonicotinoids, ◆ = organophosphates. The diagonal dashed curve indicates a decrease in application rate that is proportional to increasing LD_{50}. I.e. many of the more potent compounds are applied at higher rates than DDT, relative to their toxicity. Redrawn from Cresswell (2016).

For regulatory purposes, the limitation of the LD_{50} is that it is highly specific to the laboratory conditions that were used to determine it, which may not be environmentally realistic. For example, the conventional laboratory tests in honey bees use only healthy, newly emerged individuals whereas realistic in-hive populations comprise bees of mixed age and varied health status. Also, the LD^{50} is normally established in short-term, 'acute' exposures – typically a 48 hour exposure - which does not reveal whether toxicity could be amplified as the duration of the exposure increases. In actuality, the LD_{50} can become lower as the duration of the exposure increases (see below).

Overall, of course, the LD_{50} justifies its place among the cardinal indicators because it can be used both in setting protection thresholds and as a comparator. However, its limitations indicate that it should not be the sole cardinal number.

Second cardinal number: the NOEC

The unit of concentration specified by the NOEC (no observable effect concentration) may refer to the concentration of the toxicant in either the subject's environment or diet. Regulators can use the NOEC directly as a protection threshold, because it safeguards the focal species from obvious toxic effects. Where the test endpoint is fatality, the NOEC restricts permissible exposures to levels that do not increase the death rate above normal background levels. The NOEC does not preclude harm when used as a protection threshold, because the impact may be subtle (i.e. not observable under the examination used); only the NEC (no effect concentration) provides complete protection.

In relation to the NOEC, the term 'observable' can be taken to mean 'detectable by a specified experimental method'. Where the experimental method used to determine the NOEC is factorial (i.e. a particular number of treatments of various dosing levels are implemented), then detectability in practice means 'statistically different from the control'. Specifically, when this factorial design is used, the NOEC is taken to be the lowest of the tested doses in which the measured response of the exposed subjects is not statistically different from the response of undosed controls. Statistical tests between factor levels conventionally are based on the standard errors of the treatment means (e.g. ANOVA or t-tests), which depend on sample size because $SE = SD/\sqrt{n}$. Consequently, an undesirable situation arises where the NOEC is designated to be the smallest dose that causes a response different to the control given the size of the experiment; specifically, the value of the NOEC has no biological basis, but instead changes with the power of the experimental design. The NOEC has been criticised for this failing (Laskowski, 1995). What remedy is there? Instead of a factorial experiment, it is better in principal to characterise the dose-response relationship by curve-fitting (i.e. a regression approach) and then to estimate the NOEC

from the best fit. The question then becomes: where is the NOEC on the best-fit dose-response curve?

Insect physiologists have faced an analogous problem in estimating 'basal temperature', which is the lowest temperature at which metabolic activity begins. Their solution (Wigglesworth, 1965) has been to extrapolate from the linear section of the sigmoidal temperature-vs.-metabolic-rate relationship and to solve for an intersection with the x-axis, which is a point where metabolic rate (y-axis) is zero. It is straightforward to apply the same technique to the problem of the NOEC (Fig. 3). Here, I denote this x-intercept by NOEC* (to distinguish it from the conventional NOEC). The advantage of this approach is that experiments based on differently sized experiments will all estimate the same theorised value (the x –intercept, NOEC*) and the size of the experiment (the number of experimental subjects) affects only the confidence intervals around the estimate. In estimating NOEC*, therefore, statistical power affects only the precision of the outcome and does not bias the value of the estimator itself, unlike with the factorial/ANOVA approach described above. In an important sense, therefore, NOEC* is precisely defined and 'parametric' – it is the x -intercept of the extrapolation from the central inflection point of a sigmoidal dose-response curve.

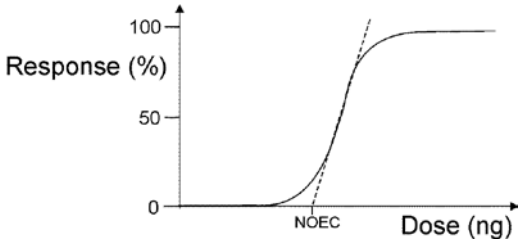

Fig. 3. A hypothetical sigmoidal dose-response relationship with a straight-line extrapolation (dashed line) to the *x*-axis from the inflection of the sigmoidal curve, which can been used to estimate the NOEC*.

A additional theoretical parameter, NEC (no effect concentration), is located where the dose-response curve leaves the abscissa (x-axis), and NOEC* is a sensible proxy for NEC under the proviso that concentrations below NOEC have an acceptably 'negligible' effect. It is an undesirable outcome that the magnitude of the so-called negligible effect is related to the gradient of the central linear section of the dose-response curve. Specifically, the extrapolation from the linear section of the dose-response curve will require slightly greater responses to be designated as 'negligible' in populations that vary more in tolerance (i.e. shallower dose-response curve; see Fig. 1). This is not entirely satisfactory and therefore we will look elsewhere for a more consistent estimator of the NOEC (see below).

Third cardinal number: Haber exponent, b.

The Haber exponent qualifies the meaning of the LD_{50} as both a comparator and a protection threshold. Its use has been recommended to toxicologists generally (Rozman, 2000) and for those interested in bee-pesticide interactions (Tennekes & Sanchez-Bayo, 2011). To discover its value, consider the impact on a regulator's decision of using the Haber exponent in conjunction with the LD_{50} to compare the hazard of two hypothetical compounds, A and B. The two compounds are intended for application as pesticides to a mass-flowering crop that blooms for several weeks. A and B have 48-hour LD_{50} values of 4 ng honey bee-1 and 2 ng bee-1, respectively. The regulator who makes a decision based on the conventional comparison between the LD_{50} values alone concludes that A and B pose a similar hazard to bees. The conventional regulator therefore approves both A and B for use provided that the application guidelines of the compounds meet satisfactory standards. Suppose, however, that the Haber exponents of compounds A and B are $bA = 1$ and $bB = 2$, respectively. The better-informed regulator who compares the LD_{50} values and takes into account the Haber values concludes (correctly) that B is much more hazardous to bees than A because the environmentally realistic exposure (several weeks) is longer than 48 hours.

The better-informed regulator safeguards farmland bees by permitting only compound A to proceed to market. What is the basis of this crucial distinction between b = 1 and b = 2?

To interpret the value of a Haber exponent, it is necessary to introduce the concept of 'toxic load'. Assume that an exposed bee is slightly injured at a constant rate by each molecule of toxicant that is inside its body. Each small injury is permanent and the bee dies when the total accumulation of injuries exceeds its individual tolerance threshold, which varies among bees (Fig 1). The toxic load is defined as the total injury accumulated by an individual bee after any particular exposure time.

The Haber exponent refers to the rate of increase of toxic load over time during an exposure; specifically, a sustained exposure to compound A (bA = 1) produces a straight-line increase in toxic load over time (Fig. 4, trajectory A) and sustained exposure compound B (bB = 2) produces a quasi-exponential increase (Fig. 4, trajectory B). (Appendix 1 presents a toxicodynamic model that relates the trajectory of toxic load to the Haber exponent.) Fig. 4 reveals an important generalisation about the relative hazard of exposure to compounds like A (b = 1) vs. B (b = 2) that have similar LD_{50} values; compounds whose Haber exponent approximates a value of b = 2 are more hazardous than compounds whose Haber exponent approximates a value of b = 1 (all else equal) provided that the case involves exposures longer than that used to determine the LD_{50} originally. In such cases, the Haber exponent is an important discriminator among toxicants.

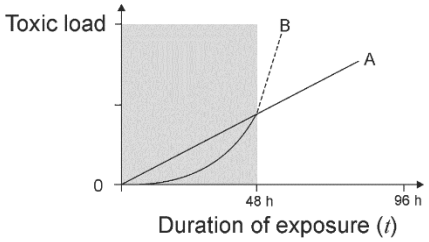

Fig. 4. Increase in toxic load (*y*-axis) over time (*x*-axis) in sustained exposures to two hypothetical compounds, A and B, that differ in the value of their Haber exponents (A: $b_A = 1$, which indicates a straight-line increase in toxic load; and B: $b_B = 2$, which indicates a quasi-exponential rate of increase).

Consequently, a regulator better safeguards bees by using the LD_{50} in conjunction with the corresponding Haber exponent.

If, as I have argued, the Haber exponent is an important qualifier of LD_{50}, how is it measured? Simply, it requires an analysis of the results of a series of 'time-to-effect' experiments, each of which is conducted at a different dose (Baas et al. 2010). A time-to-effect experiment measures the duration of exposure that is required to cause a specified effect, such as 50% mortality among exposed subjects. Typically, varied exposure levels are used in the laboratory to yield a series of 'dose-duration' combinations that cause the specific effect. For example, 20 cages each of 10 honey bees might be each exposed to one of four dietary concentrations of toxicant X (i.e. five replicates per concentration) and the investigator records the time at which the median fatality occurs in each cage. The results of the experimental series are four combinations of dose (expressed as toxicant concentration (C) and duration of exposure (t) that produce a specified effect, such as 50% mortality [i.e. (C1, t1); (C2, t2); (C3, t3); and (C4, t4)]. Normally, the duration of the required exposure, t, increases as the concentration of the toxicant, C, declines. The Haber exponent of X is evaluated by estimating the slope of the concentration-vs.-duration relationship (C-vs.-t) on a log-log plot (Appendix 1 provides a justifying explanation).

Alert readers have noticed that the preceding description makes no mention of a control treatment, which should comprise unexposed test subjects. And, in actuality, none is required in the calculation of the Haber exponent. However, a valid exponent must be estimated only from subjects under toxic exposures, because the Haber exponent is a measure of dose-dependence; dose-independent variation confounds the analysis. In order to evaluate dose-dependence, our analysis must include only C-vs.-t data recorded on subjects that the dose has detrimentally

affected. In practice, it is impossible to distinguish fatality due to the toxicant and fatality due to senescence, which can occur in individuals that tolerate the low-level doses. To exclude individuals that have not suffered fatality due to the toxicant, it is necessary to establish a statistical confidence interval on the performance of control subjects. Once this is achieved (Fig. 5), the Haber exponent can be established by regression.

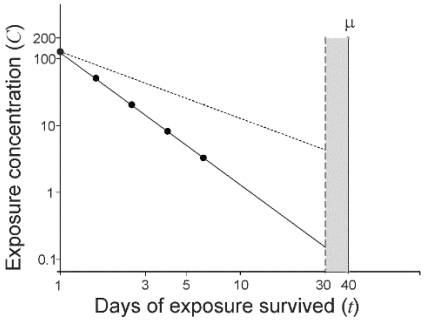

Fig. 5 An idealized *C*-vs.-*t* relationship on log-log scales for a hypothetical toxicant that causes the criterion effect (kills 50% of exposed bees in a cage) in one day when the exposure is at a concentration of 125 parts per billion (ppb). Four less concentrated exposures were tested (results also denoted by ●) and the least-squares regression has a slope of $b = -2$ (the dashed diagonal shows $b = 1$ for reference). In unexposed cages, 50% of bees died by senescence in $\mu = 40$ days and the standard deviation among cages was 5.1 days. A 95% confidence interval on the criterion in unexposed bees yields a lower boundary of 30 days (depicted by the grey-filled area). The NOEC** based on the intercept between the *C*-vs.-*t* relationship and the confidence interval is $C = 0.14$ ppb, which compares to the reference NOEC ($b = 1$) of $C = 4.2$, which is thirty times higher.

Extrapolation of the C-vs.-t relationship on log-log scales (Fig. 5) enables another estimate to be made of NOEC, which I denote NOEC** (to distinguish it from the conventional NOEC and the x-intercept estimate, NOEC*). Specifically, the intercept between the log(C)-vs.- log(t) regression and the lower confidence interval on the responses of the control population (Fig 5) is, in theory, the lowest toxic dose. It is a parametric datum whose true value is independent of sample size because the confidence interval is determined using the standard deviation (average distance of individuals from the population average), which is a population attribute (unlike the standard error, which is an attribute of the sampling procedure).

Conclusions

The Haber exponent can serve as an important qualifier of the widely used LD_{50}. The protocol used to measure the Haber exponent also enables the NOEC to be estimated. In a hypothetical but realistic example (Fig. 5), the NOEC varies by a factor of 30 depending on the magnitude of the Haber exponent, which indicates its value in evaluating the hazard a pesticide poses to bees. In future, statistical investigation will be required to establish the efficient sizes for quantifying Haber exponents.

References

Baas, J., Jager, T. & Kooijman, B. 2010. Understanding toxicity as processes in time. Science of the Total Environment, 408: 3735-3739.

Bliss, C.I. 1941. The relation between exposure time, concentration and toxicity in experiments on insecticides. Annals Of The Entomological Society Of America, 33: 721-766.

Cresswell, J.E. 2016. The impacts of agrochemical pesticides on bees in intensively cultivated farmland. *Pollination services to agriculture*. (ed. B. Gemill-Herren). The Food & Agriculture Organisation of the United Nations / Routledge, London., UK.

Laskowski, R. 1995. Some good reasons to ban the use of NOEC, LOEC and related concepts in ecotoxicology. Oikos, 73: 140-144.

Rozman, K.K. 2000. The role of time in toxicology or Haber's c×t product. Toxicology, 149: 35-42.

Tennekes, H. & Sanchez-Bayo, F. 2011. Time-dependent toxicity of neonicotinoids and other toxicants: implications for a new approach to risk assessment. Journal of Environmental and Analytical Toxicology, S4: 001.

Wigglesworth, V. B. 1965. The principles of insect physiology. Methuen, London.

Appendix 1: A toxicodynamic basis for the value of the Haber exponent

Assume that the bee is slightly injured by each molecule of toxicant that is inside its body at a constant rate. Each small injury is permanent and the bee dies when the accumulation of injuries exceeds its individual tolerance threshold, denoted T, which varies among bees (Fig 1). The toxic load, denoted Lt, is defined as the total accumulated injury after any exposure time, t, and the bee dies when Lt > T. Now consider two possible scenarios.

Scenario A: the internal concentration of toxicant at its target site, C, is constant over time. Therefore, the rate of injury is constant, Lt is proportional to the duration of the exposure, t, and we can write:

$$Lt \propto Ct \qquad \text{Eq. 1}$$

This scenario pertains when the internal concentration of a toxicant equilibrates rapidly and the biological half-life (i.e. in-body residence) of the toxicant is short relative to the total duration of the exposure, which arises if the toxicant is metabolically degraded or otherwise eliminated with rapidity.

Scenario B: the internal concentration of the toxicant at its target site increases over time as exposure continues because its biological half-life is short relative to the duration of the exposure (i.e. the toxicant bioaccumulates in the bee's body). The internal concentration is therefore a variable, denoted ϕt, whose value depends on the current duration of the exposure. Therefore, the rate of injury increases over time. If the toxicant accumulates in the bee's body at a constant rate, k1, then ϕt is given by:

$$\phi t = k1t \qquad \text{Eq. 2}$$

Under these circumstances, we can write an expression for the bee's toxic load at time t by replacing the constant C in Eq 1 by the mean value of ϕt over the time span t, which is 0.5k1t (because at the start of the exposure ϕt = 0 and at the end of the exposure it is k1t). Hence, we can write:

$$Lt \propto t2 \qquad \text{Eq. 3}$$

In summary, toxic load increases at different rates under the two scenarios. Specifically, we have:

Lt $\propto$ t (scenario A) vs. Lt $\propto$ t2 (scenario B)

In theory, therefore, the exponent takes the value b =1 if the toxicant reaches steady-state and b = 2 if the toxicant bioaccumulates.

Haber's constant product law dictates:

$$Ct^b = k \qquad \text{Eq. 4}$$

It is straightforward to evaluate b using data from a series of 'time-to-effect' experiments (Fig. 5) that quantify the exposure durations required to produce a specified level of injury in experimental subjects under various doses. The procedure involves fitting the C-vs.-t relationship and determining its slope on logarithmic axes (Bliss 1941), which estimates parameter b because the log-log version of Eq 4 is given by:

$$log(C) = -b[log(t)] + log(k) \qquad \text{Eq. 5}$$