HHS Public Access

Author manuscript

Regul Toxicol Pharmacol. Author manuscript; available in PMC 2021 February 01.

Published in final edited form as:

Regul Toxicol Pharmacol. 2020 February ; 110: 104504. doi:10.1016/j.yrtph.2019.104504.

Use of benchmark dose models in risk assessment for occupational handlers of eight pesticides used in pome fruit production.

Jane Gurnick Pouzou[1], **John Kissel**[2], **Michael G Yost**[2], **Richard A Fenske**[2], **Alison C Cullen**[2]

[1]EpiX Analytics, LLC

[2]University of Washington

1. Introduction

Despite a decades-old and often-echoed recommendation to adopt benchmark dose (BMD) modeling as the default basis for regulatory limits[1–6], the method has had slow acceptance. Use of the no observable adverse effect level (NOAEL) is still the standard procedure for derivation of a regulatory limit in many cases, including the US Environmental Protection Agency (EPA)[7] and the Organisation for Economic Cooperation and Development (OECD) Guidelines for the Testing of Chemicals[8] (for example the evaluation of neurotoxicity in rodents[9]). The limitations of the NOAEL approach have been described[3, 10]. One of the most significant is that because selection of the NOAEL depends on the identification of a statistical difference between groups, smaller group sizes are less likely to have sufficient power to identify a small effect and will therefore produce a higher NOAEL. In contrast, the BMDL method is based on uncertainty intervals – which are wider with a smaller n – and will produce a lower BMDL when group sizes reduce statistical power[11]. Furthermore, the identified point of departure necessarily must fall among the pre-selected dose groups used in the study when using the NOAEL approach. Because of this constraint, there may be very little biological basis behind the specific dose identified as a limit, despite the implications of regulatory meaning assigned and the use of the NOAEL in quantification, particularly if the study was designed in the absence of existing toxicology studies.

Travis et. al[10] identified a number of reasons why BMD methods have not been adopted with greater speed and why the NOAEL should remain the predominant tool in determining the point of departure (POD). For example, they suggest that the NOAEL is more intuitive and easier to verify and understand, that quantal outcome BMDs cannot accurately reflect the same kind of outcome in human populations since they are based on variability which may not correspond across species – also applicable to the NOAEL – and that the BMD is too sensitive to the model type selected (a criticism now addressed through model averaging methods). Despite these criticisms the BMD has become widely accepted as the more

Corresponding Author: Jane Gurnick Pouzou, Corresponding Author's Institution: EpiX Analytics, LLC

5.Declaration of conflicts of interest

The authors report no conflicts of interest. Funding in support of this analysis was provided by the NIEHS Environmental Toxicology and Pathology training grant and the CDC/NIOSH Cooperative agreement U54 OH007544.

advanced method in more recent years.[12] The intuitive appeal of the NOAEL may lead to a false sense of safety while the BMD more explicitly relies on an acceptable minimum effect of exposure, and neither the NOAEL nor the BMD is guaranteed to provide zero risk. Furthermore, the applicability of the NOAEL estimation may also suffer from differences in variability between human and animal populations[13]. Simplicity of method is not a benefit if it limits the scientific basis and usefulness of the result. The BMD method is easier to reconcile with uncertainties in dose estimates than deterministic methods such as the NOAEL, and the resulting values are more useful in probabilistic assessments of risk[14]. Wignall et al. discuss the lack of transparency and consistency in the BMD approach, and suggest that a 1 standard deviation or 10% critical effect could serve as a standard basis for unified benchmark dose modeling for large sets of toxicological data[15].

More recently, Zeller et al.[16] and Slob[17] have suggested that endpoint-specific critical effect sizes allow the most practical and most relevant assessments. Zeller and colleagues made the case that an extended standard deviation-based approach using historical control data provides the best endpoint which accounts for assay-specific and animal model-specific variability. Similarly, Slob suggests that the critical effect size for a given endpoint should be adjusted based on the maximum value of the response and the within-group variation. These measures would provide more biologically relevant points of departure; however, the Zeller et al. method requires information on the historical values of these endpoints in each animal model, which is not necessarily included in the typical study protocol.

Pesticide regulation in the United States is based on toxicological studies performed in accordance with OECD guidelines, and is one area of chemical risk assessment where the NOAEL, along with a set of uncertainty factors, is the basis of regulatory limits. Numerous tests are required for registration of an active ingredient, assessing potential human and ecological impacts of any proposed use. In the assessment of these studies, where a hazard is deemed present, one or more NOAELs is chosen to pair with residential, dietary, and occupational doses to humans based on the length of the exposure, sensitivity and specificity of the study and outcome, and the route of exposure. For (PHED) were used to construct probabilistic exposure rates similar to those deterministic rates used in official calculations. The distributions of these rates were paired with distributions of other factors, including clothing and PPE protection factors, application rates and areas, and anthropometric variables[30–32]. In the EPA human health assessments, occupational exposures to methoxyfenozide and spinetoram were calculated for the inhalation route only, as there was not considered to be evidence of a hazard via the dermal route based on acute dermal studies[27, 33]. To remain consistent with the EPA methods, probabilistic dose estimates were generated both with and without the dermal pathway for spinetoram and methoxyfenozide.

The purpose of this analysis is to demonstrate the feasibility of using OECD guideline-compliant toxicology studies carried out to generate NOAELs for production of dose-response models and benchmark doses with associated lower confidence limits for a variety of pesticides that are or were popular for the control of codling moth in tree fruit, and to determine whether the NOAEL or the BMD approach is consistently more protective against acute exposures. The effect of using a variety of critical effect sizes in continuous data on the BMD and the resulting Margin of Exposure for acute, one day doses, is also explored.

2. Methods

Eight pesticides with a variety of potential acute or sub-acute health impacts as identified by their respective EPA human health risk assessments were selected for this analysis: azinphos methyl, acetamiprid, emamectin benzoate, methoxyfenozide, novaluron, phosmet, spinetoram, and thiacloprid[20–28]. These pesticides are currently or were formerly used frequently in the production of pome fruit, and have comparable task profiles for their application The one-day dose in mg/kg body weight/day, which is the basis of occupational exposure regulation, of a pesticide handler mixing, loading, and applying each pesticide using open cab airblast methods was calculated probabilistically as described previously[29], and is also described in the supplemental material. In brief, exposure data from the Agricultural Handler Exposure Database (AHED) and the Pesticide Handler Exposure Database size. In addition, since phosmet risk assessment was based on a combination of data from an oral and dermal toxicological study, the dermal dose was compared to the points of departure from the dermal study without the inhalation dose. Second, based on the recommendation of EPA benchmark dose modeling guidelines[36], the 1 standard deviation effect size was also examined for continuous impacts. Using the method described by Slob et al.[17], the maximum value of the endpoint was used to derive a more biologically-relevant effect size for each pesticide by dividing the natural log of the maximum by eight, the Normalized Effect Size Benchmark Dose or BMD_{NES}. The hybrid method[5] was also applied using a 10% effect level for all continuous outcomes for comparison against the relative effect BMDs.

Data from the same studies used to select a NOAEL for use in EPA occupational risk assessments were used to construct multiple benchmark dose models for each pesticide. The studies, identified by Master Record Identification (MRID) in Table 1, were obtained via Freedom of Information Act request. Using the EPA Benchmark Dose Software (version 3.1.1.), available models were fit to the dataset, according to whether the endpoint was quantal or continuous. Quantal models were fit using gamma, logistic, log-logistic, log-probit, probit, Weibull, and quantal-linear models. Continuous models included exponential, Hill, linear, polynomial, and power models. The BMD Software uses maximum likelihood methods in the calculation of model parameters, and the confidence interval of the benchmark dose is calculated using profile likelihood methods[3]. A 10% effect level was used for all quantal models to permit the most comparability among models and among health outcomes. A benchmark response level was chosen for each continuous outcome in several ways to compare the result of each method.

The response or critical effect size for the continuous models was first based on levels reported by Dekkers et al. specified in a survey of experts on the commonly recommended effect size for a variety of outcomes[34]. For all endpoints except cholinesterase depression, which was assigned 20% as a known toxicologically-relevant effect size[35], this level was 10% relative deviation from the control group. Phosmet and azinphos methyl were modeled with both 10% and 20% depression as the critical effect occupational exposures the NOAEL is divided by an estimate of the human dose over a single day to calculate the Margin of Exposure (MOE)[18], which must be above a Level of Concern (LOC) – 100 being the most commonly used[7]. Substitution of a benchmark dose into this existing paradigm should be

feasible if the existing studies can be used to generate dose-response curves. A further advantage of this method is that integration of new approaches involving physiologically-based pharmacokinetic (PBPK) model development can be integrated with the dose-response models used in production of the BMD[19].

2.1. Selection of toxicological outcomes for benchmark dose models

2.1.1. Acetamiprid—Symptoms of developmental neurotoxicity were recorded in rats through a functional observation battery of neurotoxicity testing in offspring of orally-dosed dams (for 6 weeks perinatally). The most relevant acetamiprid outcome recorded was change in the auditory startle reflex amplitude maximum in males at post-natal days (PND) 20 and 60. Other outcomes (including reductions in pup viability and alterations in weight gain) were non-specific to neurodevelopmental impacts.

2.1.2. Azinphos methyl—The outcome of interest for assessment of azinphos methyl neurotoxicity was identified *a priori* as cholinesterase depression. Erythrocyte, plasma, and brain cholinesterase were measured at varying time points during the 1-year feeding study performed with dogs, and at 13 weeks, significant depressions in all three cholinesterases' activity were recorded. Although acute studies assessing cholinesterase inhibition in animals and in humans were available, this study was judged by the EPA to be the most protective and appropriately conducted for comparison with biomonitoring dose measurements of pesticide applicators.

2.1.3. Emamectin benzoate—The emamectin benzoate feeding study of acute (15-day) neurotoxicity in mice was used to find the most protective NOAEL for occupational risk assessment. Several endpoints indicative of neurotoxicity were recorded; tremors were the first frank symptom to be observed, followed by ptosis, gait and posture abnormalities, decreased activity, urine staining, and labored breathing. At necropsy, some animals in the highest dose groups had sciatic nerve degeneration. Since tremors appeared first and at the lowest doses of all symptoms, they were regarded as the most sensitive indicator and selected for this analysis and by the EPA for derivation of the NOAEL.

2.1.4. Methoxyfenozide—Occupational exposures to methoxyfenozide were assessed by the EPA only for inhalation, as acute dermal toxicity studies did not indicate a hazard according to the EPA human health risk assessment for this compound. Although various other outcomes were investigated, hematological impacts shown in a two-week feeding study of dogs were selected to derive the NOAEL used in the occupational risk assessment. This study included only two animals per sex per dose group; however, similar hematological toxicity was observed at 3 months in the 1-year study with a sample size of 4 per group. This study was therefore used to derive the benchmark dose model.

2.1.5. Novaluron—The NOAEL used for occupational risk assessment of Novaluron handlers was drawn from a 90-day feeding study performed in rats which assessed a variety of hematologic parameters. Red blood count, hematocrit, and hemoglobin were all influenced in the higher dose groups. In addition, spleen and liver pigmentation and splenic

hematopoiesis were observed, and a combination of all of these impacts led to the derivation of the NOAEL from this study.

2.1.6. Phosmet—Like azinphos methyl, phosmet is known to act as a cholinesterase inhibitor, therefore the outcome of cholinesterase activity was measured in acute toxicity studies. In the case of phosmet, a dermal and oral study were both used to generate separate NOAELs for the routes of exposure, since there was no human biomonitoring data as with the case of azinphos methyl.

2.1.7 Spinetoram—As with methoxyfenozide, insufficient evidence of dermal toxicity in the short or intermediate term was found in the EPA human health risk assessment to warrant a complete risk assessment of that exposure route beyond the hazard assessment stage. The outcomes used in the derivation of the NOAEL for the inhalation route of exposure were also hematologic, drawn from the sub-chronic (90-day) feeding study performed with dogs. Blood cell and hematocrit levels were significantly affected, and anemia, arteritis and bone marrow necrosis were observed. The anemia and lowered platelet counts observed were believed to be secondary to the bone marrow necrosis, so the arteritis and bone marrow necrosis were both evaluated.

2.1.8. Thiacloprid—The normal battery of toxicological evaluations showed a number of potential impacts of thiacloprid dosing. Occupational doses were evaluated by the EPA using a NOAEL derived from liver and thyroid impacts observed in the subchronic inhalation and chronic feeding studies of rats. A variety of liver impacts were recorded, including enzymatic induction (N-demethylase, O-demethylase, and CYP450), and hepatocellular hypertrophy. Thyroid hypertrophy was also noted. Of the two organs, the liver impacts were evaluated at lower doses. Thiacloprid is also classified as a likely human carcinogen, but as the EPA occupational risk assessment was based on organ toxicities as the more protective outcomes, a cancer risk assessment was not performed in this analysis.

2.2 Model assessment

Each possible model was assessed for goodness-of-fit using qualitative evaluation of the dose-response graph and the p-value of the X^2 goodness-of-fit test. Models were compared within each toxicological endpoint using the qualitative fit of the curve, the Akaike Information Criterion (AIC), and the residuals. The 95% confidence limit was calculated and the resulting BMD and the lower confidence limit of the benchmark dose (BMDL) were compared with the NOAEL from the same study and the deterministic dose used in the EPA human health assessment for the same pesticide application scenario. The estimated dose distribution was also compared with all three points of departure (BMD, BMDL, and NOAEL) using the calculated exceedance fraction, defined as the proportion of the dose distribution which is above the POD, computed using the *efraction.exact* command of the R package STAND[37].

3. Results

The toxicological studies and endpoints of interest for each pesticide are described in Table 1. Although each compound was associated with multiple studies which yielded

toxicological endpoints potentially useful in creating a benchmark dose, the studies presented here are those used in the EPA's human health risk assessment for pesticide handlers to explore the impact of use of a benchmark dose in place of a NOAEL. Goodness of fit values and the resulting BMD and BMDL for each possible model fit to the eight pesticides' outcomes can be found in Supplemental Table B. The graphs for the selected models are shown in Figure 1. Parameters for the chosen models are listed in supplemental material.

3.1. Selection of toxicological outcomes and benchmark dose models

3.1.1. Acetamiprid—The outcomes not specific to neurodevelopmental toxicity observed in this study were not able to produce a dose-response model due to variability in the control animals. PND 20 was selected for this analysis as the dose-response effect was more evident than at PND 60. Of the models assessed, similar results for goodness of fit and residuals were generated, and the Hill model had the lowest AIC.

3.1.2. Azinphos methyl—The erythrocyte cholinesterase outcome of the 1-year study provided the most protective result and an advantage over the brain cholinesterase measurement in that it had been checked at 4 weeks after baseline as well. This 4-week time point in males was used as the basis for the dose-response as it was closer to the length of exposure expected in an occupational setting. However, the results from the 13-week measurement produced a lower benchmark dose (at 10% effect size, 0.23 mg/kg/day compared with the 13-week study's 0.07 mg/kg/day), although the same NOAEL would be selected based on either time point.

3.1.3. Emamectin benzoate—All dichotomous models were successfully fit to the dose-response of tremors in the 15-day study, passing goodness of fit testing with satisfactorily low residuals (considered to be residuals less than |2.0|). The AIC values were similar, with the lowest being the quantal-linear model. This model also was the best fit after visual assessment of the curve.

3.1.4. Methoxyfenozide—The authors of the identified 2-week study noted symptoms but did not believe them to be treatment related. However, dose-responsive patterns were found in the male treatment groups. The outcome for which model fitting was successful was the three-month measurement of red blood cell count, in contrast with platelet count, red blood cell count, hematocrit, and methemoglobin. All continuous models showed satisfactory fit, and the exponential 4 model was selected based on a marginally lower AIC.

3.1.5. Novaluron—With the exception of red blood cell counts and hematopoiesis, the dose-response for hematological impacts was irregular and non-significant, resulting in poor model fits. The outcome of reduced blood cell count was therefore used in the calculation of the benchmark dose. Models for both the red blood cell count and spleen hematopoiesis passed the goodness of fit and variance tests and had comparable AIC among them. Either endpoint could be used, but the dose-response is clearer in the RBC data. The exponential 4 model of RBC was selected based on AIC and visual evaluation of the model fit.

3.1.6. Phosmet—Dose-response curves were constructed using both dermal and acute oral data, for plasma ChE in the dermal study and red blood cell cholinesterase in the oral study, which proved to be the most sensitive measures. For both the dermal and oral models, the Hill model provided the best fit and lowest AIC. The model for the oral exposure was a better fit based on the results of the χ^2 goodness-of-fit test as well as providing a lower benchmark value, as expected.

3.1.7. Spinetoram—Since all animals in the control and lowest dose group were free of arteritis and all animals in the higher dose groups developed it, the dose-responses for arteritis were of limited use, therefore the bone marrow necrosis was used to derive the benchmark dose. All dose-response curves for necrosis were similar overall with respect to AIC and residual values, but the logistic model offered the best fit qualitatively for both the dose-response curve and its 95% confidence interval.

3.1.8. Thiacloprid—The liver enzyme induction impacts in general produced model fits which identified significant dose responses and passed goodness-of-fit tests for the mean, but in many cases the dose-response was inconsistent in the low-dose groups leading to a poorer model fit, particularly at the low doses. The model which was selected was the log-logistic model of hepatocellular hypertrophy, which provided a more consistent fit to the data at low doses (based on visual inspection of the curve), passed the goodness-of-fit test, and had satisfactory residual values.

3.2 Benchmark doses and NOAELs in comparison to deterministic dose

The ratio of the 10% benchmark dose to the associated lower 95% confidence interval ranged from 1.02 to 5.76, with a mean of 3.3 (Table 2). The same ratio for the BMD_{NES} ranged from 1.58–9.6, with a mean of 3.6. Three of the benchmark doses of either type were lower than the NOAEL from the same study, and five of the $BMDL_{10}$ were lower than the NOAEL (Figure 2), whereas all of the $BMDL_{NES}$ were lower than the NOAEL except phosmet which was similar but slightly higher. The NOAELs for acetamiprid, novaluron, and phosmet were above the BMD for the same study, indicating that in those cases the NOAEL is less sensitive than the BMD method. The NOAEL for azinphos methyl, spinetoram, and emamectin benzoate fell below the BMDL, so that in those cases, the NOAEL was more protective than the BMD. The BMD for emamectin benzoate falls above the range of the data used to derive the model, and should therefore be interpreted with caution, as it represents an extrapolation of the dose response curve which may or may not be supportable. In the remaining cases of methoxyfenozide and thiacloprid, the NOAEL fell between the BMD and BMDL. The ratio of the 10% BMD to the NOAEL ranged from 0.17 to 12.15 and averaged 2.8 (Table 3), and the BMD_{NES} to NOAEL ratio averaged 1.96 (ranging from 1.04 to 4.22). The Normalized Effect Sizes ranged from 20.75% to 30.86%, and so the associated BMDs were all higher than the 10% effect size BMDs, as were the BMDs from the hybirid risk model.

The comparison of the NOAEL with the occupational handler doses used in the EPA risk assessments using an MOE supports the findings of the EPA assessments (Table 4). The only ratio which falls below the LOC, or level of concern (100 for all occupational doses except

emamectin benzoate, which has an LOC of 300), is that of azinphos methyl, as well as phosmet depending on the method of calculation and data source. For the most part, the ratio of the BMD to the dose produces the same conclusion, except in the case of acetamiprid. If the BMDL were used, azinphos methyl and acetamiprid would both produce a ratio less than 100 and therefore of concern, but all other chemical exposures would still be on average below the level of concern. The alternative of single standard deviation as the effect size for the continuous effects found a higher benchmark dose in all cases except phosmet.

3.3 Probabilistic dose comparisons with BMD and NOAEL

Comparison of the probabilistic dose estimations with the various points of departure allows estimation of an exceedance fraction, the proportion of the estimated potential doses which are above the deterministic level of concern, and the margin of exposure (MOE) calculated by dividing the BMD measures or the NOAEL by the dose. Table 4 shows this fraction for the NOAEL, BMD, and BMDL point estimates divided by 100 (300 in the case of emamectin benzoate) to account for uncertainty factors, and the MOE as well as the amount exceeding the level of concern (100 or 300 for emamectin benzoate). Exceedance fractions and MOE varied, depending on chemical and point of departure selected. Exceedance of the NOAEL ranged from 0.005 % in the case of methoxyfenozide to 72.2% for azinphos methyl. Azinphos methyl doses had the highest fraction exceeding all of the points of departure, and by extension, the highest fraction of MOE beyond the LOC. Exceedance fractions for methoxyfenozide and spinetoram were increased by a minimum of 40 times by the addition of the estimated dermal doses. The exceedance fractions for azinphos methyl and phosmet varied depending on the choice of point of departure and the source of the dose-response data (oral vs dermal). The ratio of the exceedance fractions of the BMDL and the BMD are a crude relative measure of the uncertainty in the benchmark dose measurement. These ratios ranged from 1.1 to 8.4. By comparison, the ratio of the exceedance fraction of the BMD to that of the NOAEL ranged from 0.1 to 5.0, and averaged at 1.2, indicating the relative protective ability of the two points of departure. There was no correlation between the two ratios.

1. Discussion

This analysis shows that studies designed for the production of a NOAEL according to OECD guidelines can be used to generate a dose-response curve and derive a benchmark dose and the associated confidence interval, but that the existing standards could be built upon to improve the quality of data for benchmark modeling. With the exception of azinphos methyl, there are not existing benchmark dose models available for comparison of results. In the case of azinphos methyl, the 20% BMD and BMDL in male rats found here (0.50 mg/kg/day and 0.35 mg/kg/day) are similar to the values reported by the ATSDR benchmark analysis (0.48 and 0.30 mg/kg/day), which used the same data, although different models were selected[38]. This analysis also showed that benchmark dose methods do not produce inherently more (or less) conservative or protective dose limits than the NOAEL method, as the NOAEL for many of the compounds was lower than the BMD, and in some cases less than the lower confidence limit. In a few cases, the NOAEL is within the confidence limit of the BMD or BMDL, which may indicate that they would produce indistinguishable results

for regulation. However, the BMD and BMDL still provide more information on the uncertainty of the outcome.

The relationship between the BMD_{SD}, the BMD_{NES} and the other points of departure is instructive as well, as these alternative effect sizes represent the variability in continuous outcomes and therefore the degree to which the NOAEL or BMD represents a change beyond that variability. The MOE results from the BMD-SD of Methoxyfenozide, for example, is similar to the result of the BMD_{10}, but in the case of Phosmet, BMD-SD is closest to the NOAEL. In cases such as Azinphos methyl, Acetamiprid, and Novaluron, the BMD-SD is higher than any other point of departure, which may suggest that the 10% and 20% effects are within the outcome's variability. The BMD_{NES} effect size of 26% for azinphos methyl and phosmet is close to the 20% value typically used as biologically relevant for cholinesterase inhibitors. In all cases, however, the NES was higher than the standard effect sizes. This result might be interpreted to suggest that the endpoints have too much variability to support use of a standard effect size as low as 10%, but the NES is based on a scaling factor of 1/8, which is itself an arbitrary choice, even though the overall method connects the effect size to the variation in the endpoint. A larger scaling factor, for example 20, would produce an NES of closer to 10% in most cases.

The degree of uncertainty in the BMD estimates shown here varies, and in some instances health outcomes were observed in the study which could not be used to build a dose-response model, either due to infrequent observation, lack of significant dose-response, or, non-monotonic or inconsistent dose-response. Although a successful dose-response model was fit in the case of each pesticide in this study, it is likely that in other instances, particularly where the compound is of relatively low toxicity so that responses will not be measured at lower doses, the OECD guidelines will not produce a study with sufficient data to create a model, in which case the NOAEL could serve as a contingency method rather than requiring an additional study. In several of the modeling attempts, some outcomes did not yield a model fit and so were passed over in favor of others. This effect could bias the choice of outcome towards those with less steep dose-responses, since the shorter curve would be less likely to show impacts between the minimum and maximum at doses that are intermediate for endpoints with a shallower curve. Quantal outcomes may be less likely to be selected for the same reason, especially where group sizes are smaller. An increase in the standard required number of dose groups required could help produce studies more amenable to a benchmark dose analysis and reduce the selection of shallower dose-responses, as suggested by Slob in 2002[39].

It is important to recognize that a NOAEL is as likely as the BMD to be unreliable or impossible to determine in cases where the dose-response is uncertain or variability in the controls is high[40]. The key advantage of dose-response modeling and the benchmark dose, and a compelling reason to adopt these methods as the status quo, is that some measure of the point of departure's uncertainty is available and expressed relatively simply as a confidence interval, whereas the uncertainty in estimation of a NOAEL is potentially the same but left opaque if the number is taken at face value. The second advantage to the BMD method in regulation is the incorporation of the entire dose response curve, which leaves the outcome less vulnerable to biases introduced by the dose selection and study design. A

further advantage of the dose response method is also illustrated through the comparison of the results of the cholinesterase inhibition models. The use of a continuous endpoint in the NOAEL paradigm may require only a statistical difference in the outcome to determine the target value. No biological justification or clinical significance is necessarily required. While this method may arguably be more sensitive to small changes in a measure, it is not guaranteed to be more protective, and does not necessarily lead to a result which is useful in risk management. Since the benchmark dose method requires that an effect size be specified, the model is more flexible and as demonstrated here, can be scaled to the variability of the endpoint of interest. In the case of azinphos methyl, use of 20% rather than 10% inhibition of cholinesterase decreases the proportion of the workers predicted to receive doses over the level of concern from 80% to 56%. However, for phosmet, the same decision changes the proportion from 14% to 9%. The impact of this type of difference on a risk management decision is not clear, but the potential for evaluating the sensitivity of the population's level of concern to the effect size chosen has great value in increasing the flexibility and transparency of risk management. Potentially, greater data availability on the assessment of dichotomous outcomes and the effective critical effect size decision-making they imply (as described by Slob and Pieter[14]) could allow for sensitivity analysis for the designation of outcomes usually considered quantal, for example, cellular hypertrophy.

The potential dependence of the benchmark dose on the choice of critical effect size has been described as a weakness of the method[10]. This study demonstrates that while effect size may have a large or small impact on the benchmark dose, consistency of reporting the process and results of benchmark dose modeling at these different possible effect sizes can provide transparency while quantifying uncertainty in the data. Biological basis, natural variation in the endpoint for the animal model, and transparency in the choice of critical effect size may be a more sound basis for developing a consistent BMD methodology than choice of effect size as the basis for consistency, but these methods too may suffer from arbitrariness in the selection of the cut-off value.

The methods used by the EPA to evaluate occupational exposures are based on single-day estimates for non-carcinogenic outcomes, and therefore the dose estimates used in this analysis are also for one day. However, in cases where doses may not clear from the body within 24 hours, or may have longer-term sequelae, a longer period of exposure may be important even in occupational scenarios where the assumption is that exposures are acute. The estimation of the dose over a longer period of time requires additional data and modeling of the pharmacodynamics of the compound of interest as well as data on the task patterns of applicators working for multiple days in a row using the same compound or compounds with similar toxicological impacts, which likely is highly variable.

In summary, this analysis demonstrates the use of existing OECD guideline studies to build benchmark dose models and derive points of departure for risk assessment. The use of the benchmark dose compared to the NOAEL may or may not substantively impact the risk assessment outcome but is able to provide a quantification of the uncertainty around the selected point of departure which is absent in the reports of NOAELs from the same studies. Benchmark doses provide transparency and flexibility and can be performed with existing study guidelines, despite room for improvement in study design. Consistency in the process

of modeling and reporting can provide the standardization necessary for the adoption of these measures into standard operating procedures for official risk assessment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding Body Information

The authors report no conflicts of interest. Funding in support of this analysis was provided by the NIEHS Environmental Toxicology and Pathology training grant T32ES007032 and the CDC/NIOSH Cooperative agreement U54 OH007544. The funding organizations had no influence on the planning, data collection, analysis, or manuscript preparation.

References

1. Filipsson AF, Sand S, Nilsson J, Victorin K. The benchmark dose method--review of available models, and recommendations for application in health risk assessment. Crit Rev Toxicol 2003;33(5):505–42. [PubMed: 14594105]
2. Gaylor D, Ryan L, Krewski D, Zhu Y. Procedures for calculating benchmark doses for health risk assessment. Regulatory Toxicology and Pharmacology 1998;28(2):150 164. [PubMed: 9927564]
3. Davis JA, Gift JS, Zhao QJ. Introduction to benchmark dose methods and U.S. EPA's benchmark dose software (BMDS) version 2.1.1. Toxicol Appl Pharmacol 2011 7 15;254(2):181–91. [PubMed: 21034758]
4. Crump K A new method for determining allowable daily intakes*1. Fundamental and Applied Toxicology 1984;4(5):854, 871. [PubMed: 6510615]
5. U.S. EPA. Benchmark dose technical guidance. Washington, D.C: U. S. Environmental Protection Agency; 2012 June 2012 Report nr EPA/100/R-12/001.
6. EFSA. Guidance of the scientific committee on use of the benchmark dose approach in risk assessment. EFSA Journal 2017 24 January, 2017;15(1):4658.
7. U.S. EPA. Revised risk assessment methods for workers, children of workers in agricultural fields, and pesticides with no food uses. Washington, D.C20460: Office of Pesticide Programs U.S. Environmental Protection Agency; 2009 December 7, 2014. Report nr EPA-HQ-OPP-2009-0889-0002.
8. OECD. Organisation for economic cooperation and development guidelines for the testing of chemicals, section 4. 2018.
9. OECD. Test no. 424: Neurotoxicity study in rodents In: OECD guidelines for the testing of chemicals, section 4. Paris: OECD Publishing; 1997. .
10. Travis KZ, Pate I, Welsh ZK. The role of the benchmark dose in a regulatory context. Regul Toxicol Pharmacol 2005 12;43(3):280–91. [PubMed: 16143439]
11. Oberg M Benchmark dose approaches in chemical health risk assessment in relation to number and distress of laboratory animals. Regul Toxicol Pharmacol 2010 December 2010;58(3):451. [PubMed: 20800084]
12. Benford D, Halldorsson T, Hardy A, Jeger MJ, Knutsen KH, More S, Mortensen A, Naegeli H, Noteborn H, Ockleford C, et al. Update: Use of the benchmark dose approach in risk assessment. EFSA Journal 2017 24 January 2017;15(1).
13. Allen B Dose-response assessment for developmental toxicity II. comparison of generic benchmark dose estimates with no observed adverse effect levels. Fundamental and Applied Toxicology 1994;23(4):487 495. [PubMed: 7867900]
14. Slob W, Pieters MN. A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: General framework. Risk Analysis 1998;18(6):787 798. [PubMed: 9972582]

15. Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, Rusyn I. Standardizing benchmark dose calculations to improve science-based decisions in human health assessments. Environ Health Perspect 2014 5;122(5):499–505. [PubMed: 24569956]
16. Zeller A, Duran-Pacheco G, Guerard M. An appraisal of critical effect sizes for the benchmark dose approach to assess dose-response relationships in genetic toxicology. . Arch Toxicol 2017 Dec 2017;91(12):3799–807. [PubMed: 28799093]
17. Slob W A general theory of effect size, and its consequences for defining the benchmark response (BMR) for continuous endpoints. Crit Rev Toxicol 2017 Apr 2017;47(4):342–51. [PubMed: 27805866]
18. OECD. WHO OECD ILSI/HESI international workshop on risk assessment of combined exposures to multiple chemicals. Paris: OECD; 2011 February 15, 2011. Report nr Series on Testing & Assessment No. 140.
19. Guidance for Applying Quantitative Data to Develop Data-Derived Extrapolation Factors for Interspecies and Intraspecies Extrapolation [Internet]: U.S. EPA; c2015 [cited 2016 6/20/2016]. Available from: https://www.epa.gov/risk/guidance-applying-quantitative-data-develop-data-derived-extrapolation-factors-interspecies-and.
20. US EPA. Emamectin benzoate human health assessment scoping document in support of registration review. Office of Chemical Safety and Pollution Prevention: US Environmental Protection Agency; 2011 June 16 2011:Available at: https://www.regulations.gov/document?D=EPA-HQ-OPP-2011-0483-0004.
21. US EPA. Human health risk assessment for the section 3 registration of acetamiprid on cotton, leafy vegetables, brassica vegetables, fruiting vegetables, citrus, pome fruits, grapes, and canola and mustard seed. US Environmental Protection Agency; 2005 Report nr EPA-HQ-OPP-2005-0190-0011:Available at: https://www.regulations.gov/document?D=EPA-HQ-OPP-2005-0190-0011.
22. US EPA. Revised occupational exposure and risk assessment for azinphos methyl (reflecting recommendations from the human studies review board). US Environmental Protection Agency Office of Prevention, Pesticides, and Toxic Substances; 2006 Report nr EPA-HQ-OPP-2005-0061-0136.
23. US EPA. Novaluron: Human health risk assessment for proposed section 3 uses on bushberry crop subgroup 13–07B; brassica, leafy greens, crop subgroup 5B; turnip, greens; and fruit, stone crop group 12. Office of Prevention, Pesticides, and Toxic Substances: US Environmental Protection Agency; 2009 Sept 9, 2009.
24. US EPA. Methoxyfenozide human health risk assessment to support proposed new uses on herbs, caneberries, dates and sorghum; to establish rotational crop tolerances in the rapeseed and sunflower oilseed subgroups; as well as to extend and update crop group tolerances on multiple commodities. Office of Chemical Safety and Pollution Protection: US Environmental Protection Agency; 2013 August 6, 2013.
25. US EPA. Phosmet (chemical ID 059201) HED human health risk assessment and supporting documentation for the reregistration eligibility decision document (RED). Office of Prevention, Pesticides, and Toxic Substances: US Environmental Protection Agency; 1998 October 30, 1998.
26. US EPA. Pyriproxyfen: Human health risk assessment for the request to add uses on herb subgroup 19A, and the expansions of existing crop group uses to numerous crop subgroups. Office of Chemical Safety and Pollution Prevention: US Environmental Protection Agency; 2011 September 8, 2011.
27. U.S. EPA. Spinosad and spinetoram: Human-health assessment scoping document in support of registration review. US Environmental Protection Agency 2011 August 18, 2011.
28. US EPA Thiacloprid in/on pome fruits and cotton health effects division risk assessment. US Environmental Protection Agency Office of Prevention, Pesticides, and Toxic Substances 2003 July 23, 2003.
29. Pouzou JG, Cullen AC, Yost MG, Kissel JC, Fenske RA. Comparative probabilistic assessment of occupational pesticide exposures based on regulatory assessments: Comparative probabilistic assessment of occupational pesticide exposures. Risk Analysis 2017 November 8, 2017.
30. AHETF. Agricultural handlers exposure task force (AHETF) volume IV: Standard operating procedures. In: ; April 7, 2008. .

31. US EPA. U.S. EPA. exposure factors handbook 2011 edition (final). Washington, DC: US Environmental Protection Agency Office of Research and Development; 2011 Report nr EPA/600/R-09/052F.
32. U.S. EPA. Development of statistical distributions or ranges of standard factors used in exposure assessments. Exposure Assessment Group, Office of Health and Environmental Assessment, US Environmental Protection Agency 1985 August, 1985;PB85–242667.
33. U.S. EPA. Pyriproxyfen human health risk assessment for the request to add uses on herb subgroup A, and the expansions of exisiting crop group uses to numerous crop subgroups. Office of Chemical Safety and Pollution Prevention; 2012 August 22, 2012.
34. Dekkers S, de Heer C, Rennen MAJ. Critical effect sizes in toxicological risk assessment: A comprehensive and critical evaluation. Environ Toxicol Pharmacol 2001;10(1–2):33 52. [PubMed: 11382555]
35. Buist HE, von Bölcsházy G, Dammann M, .Telman J, Rennen MAJ. Derivation of the minimal magnitude of the critical effect size for continuous toxicological parameters from within-animal variation in control group data. Reg Toxicol Pharmacol 2009 Nov 2009;55(2):139–50.
36. US EPA. Benchmark dose technical guidance. Risk Assessment Forum U.S. Environmental Protection Agency; 2012 June 2012. Report nr EPA/100/R-12/001.
37. Frome EL, Frome DP. STAND: Statistical analysis of non-detects. R package version 2.0. http://CRAN.R-project.org/package=STAND. 2015.
38. ATSDR. Toxicological profile for guthion. U.S. department of health and human services, public health service, agency for toxic substances and disease registry. 2008 September 2008;tp188.
39. Slob W Dose-response modeling of continuous endpoints. Toxicol Sci 2002 4;66(2):298–312. [PubMed: 11896297]
40. WHO Task Group on Environmental Health Criteria on Principles for Modelling Dose-Response for the Risk Assessment of Chemicals., United Nations Environment Programme., International Labour Organisation.,Inter-Organization Programme for the Sound Management of Chemicals.,. Principles for modelling dose-response for the risk assessment of chemicals. Geneva, Switzerland: World Health Organization; 2009 ID: 264018462.
41. Nemec M Acetamiprid: An oral developmental neurotoxicity study in rats. laboratory project ID WIL-21193. Ashland, Ohio: WIL Research Laboratories, Inc.,; 2003 November 21, 2003. Report nr 462556–19.
42. Allen TR. 52-week oral toxicity (feeding) study with azinphos-methyl (E 1582) in the dog. Itingen, Switzerland: Mobay Corporation (Bayer AG); 1990 May 31, 1990. Report nr 100644.
43. Gerson RJ. L-660,599: Fifteen day dietary neurotoxicity study in CF-1 mice. MRID 428515–03. Merck & Co.; 1993 April 2, 1993. Report nr 92-049-0.
44. Morrison RD, Shuey DL. RH-2485 technical: One-year dietary toxicity study in dogs. Report 94R-257. Spring House, PA: Rohm and Haas Company; 1997 May 21, 1997. Report nr 446177–28.
45. Kirk SJ. GR 572 (technical) toxicity to rats by dietary administration for 13 weeks. Lab. project no.: AGR 50/90386. Cambridgeshire, England: Makhteshim Chemical Works Ltd.; 1990 July 2, 1990. Report nr 45615–03.
46. Cappon GD. An acute neurotoxicity study of phosmet in rats. WIL-331004. Yuma, AZ: Gowan Company; 1998 October 8, 1998. Report nr 44673301.
47. Hilaski RJ. A 21-day dermal toxicity study of imidan in rats. Mattawan, MI: Gowan Company; 1999 March 18, 1999. Report nr 447958–01.
48. Stebbins KE, Brooks KJ. XDE-175: 90-day dietary toxicity study in beagle dogs. Midland, MI: Dow AgroSciences, LLC; 2005 May 31, 2005. Report nr 465685–01.
49. Pauluhn J YRC 2894: Subacute inhalation toxicity on rats (exposure 5 × 6 hours/week for 4 weeks). Wuppertal, Germany: Bayer AG; 1998 June 17, 1998. Report nr 449211–15.

Highlights

1. Benchmark doses were estimated for eight pesticides using toxicological studies completed for registration.
2. The Benchmark doses/confidence limits were compared with the NOAEL as the regulatory POD; none were consistently lowest.
3. Occupational doses were compared with each POD as MOEs with exceedance fractions to compare the protection provided.
4. Current guidelines may produce BMDs as well as NOAELs, with the advantage of quantifying the uncertainty about the POD.

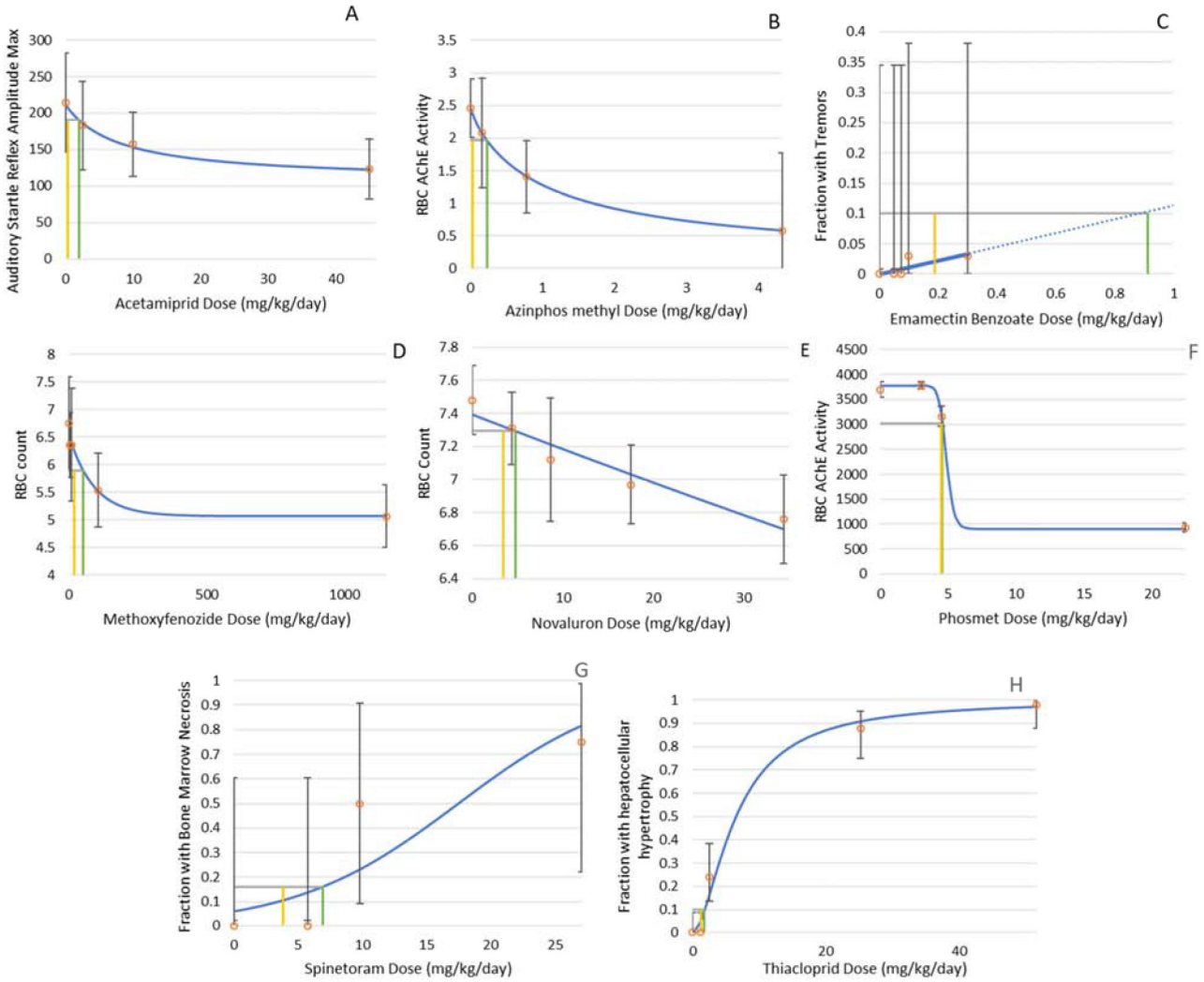

Figure 1:

Graphs of dose-response modls for the selected outcome for each pesticide. The dashed line represents the benchmark dose and confidence interval associated with an alternative critical effect size of 1 standard deviation from the control. The solid line benchmark dose is associated with the selected critical effect size listed in table 2 A) Hill model for acetamiprid-induced decreased maximum amplitude of auditory startle B) Exponential model of erythrocyte acetylcholinesterase activity and oral dose of azinphos methyl C) Quantal-linear model of emamectin benzoate-induced tremors D) Exponential model of decreased red blood cell count associated with methoxyfenozide dosing E) Exponential model of decreased red blood cell count associated with Novaluron dosing F) Hill model of erythrocyte acetylcholinesterase activity and oral dose of phosmet G) Logistic model of fraction of population with bone-marrow necrosis induced with spinetoram dosing H) Log-logistic model of hepatocellular hypertrophy associated with thiacloprid dosing

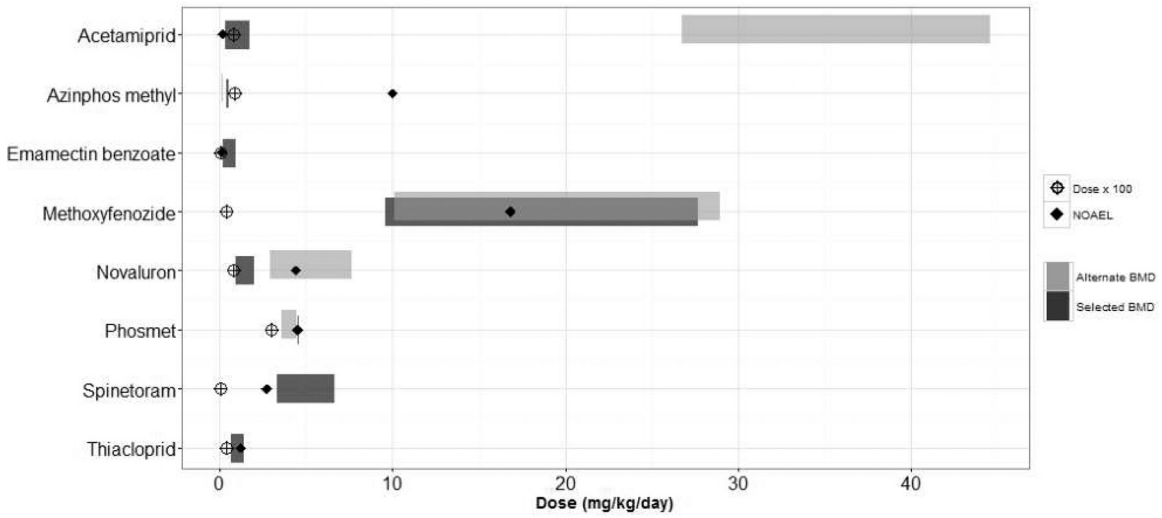

Figure 2:

NOAEL, EPA-calculated daily dose x 100 of active ingredient to a mixer/loader/applicator in pome fruit using open cab application as published in the EPA HHR risk assessments, and BMDL-BMD range for selected critical effect size, and alternative effect size for continuous endpoints. Selected effect size was 10% for all quantal impacts, 10% for all continuous except azinphos methyl and phosmet, and 20% for azinphos methyl and phosmet (cholinesterase inhibitors). The alternate effect size is 1 standard deviation from baseline. The dose x 100 represents the EPA-estimated human dose combined with the uncertainty factors used to adjust the points of departure from animal studies for comparison. The dose x 100 therefore represents the minimum toxicological POD that would produce an acceptable occupational margin of exposure under the current regulatory system in the United States, based on the doses predicted by the EPA to occur in one day of work.

Table 1:

Study MRID and selected toxicological endpoints for each pesticide. RBC = red blood cell. The original studies which have been made public through the Freedom of Information Act request may be accessed through the EPA Office of Pesticide Programs Chemical Search at https://iaspub.epa.gov/apex/pesticides/f?p=chemicalsearch:1

Pesticide	Study MRID	Endpoint Used in BMD Model	Timing of Dose	Class of Outcome	Type	Species	Dose Route
Acetamiprid	462556–19[41]	Changes in auditory startle	Perinatal (6 week)	Neurodevelopmental	Continuous	Crl:CD(SD) IGS BR rat	Oral gavage of the dams
Azinphos methyl	100644[42]	RBC Cholinesterase	Subchronic (13 week)	Neurotoxicity	Continuous	Beagle	Mixed with food
Emamectin benzoate	428515–04[43]	Tremors	Acute (2 week)	Neurotoxicity	Quantal	Crl:CF-1 BR mouse	Mixed with food
Methoxyfenozide	446177–28[44]	RBC count	Subchronic (10 week)	Hemotoxicity	Continuous	Beagle	Mixed with food
Novaluron	456515–03[45]	RBC count	Subchronic (13 week)	Hemotoxicity	Continuous	Crl:CD(SD) BR rat	Mixed with food
Phosmet	446733–01[46]	RBC Cholinesterase	Acute (one dose)	Neurotoxicity	Continuous	Crl:CD(SD) IGS BR rat	Oral gavage
Phosmet (dermal)	447968–01[47]	RBC Cholinesterase	Acute (3 week)	Neurotoxicity	Continuous	Crl:CD(SD) BR rat	Dermal (in water)
Spinetoram	465685–01[48]	Bone marrow necrosis	Subchronic (13 week)	Hemotoxicity	Quantal	Beagle	Mixed with food
Thiacloprid	449277–15[49]	Hepatocellular hypertrophy	Subchronic inhalation (13 week)	Hepatotoxicity	Quantal	Hsd Cpb: WU (SPF) Rat	Directed flow noseonly aerosol

Table 2:

Selected critical effect size, NOAEL from the investigated study, and Benchmark Dose in mg/kg/day with 95% Confidence limit for the critical effect size (CES) and alternate effect sizes (1 standard deviation for all continuous outcomes, and 10% inhibition for cholinesterase inhibitors all in mg/kg/day). The EPA-calculated dose in mg/kg/day for pesticide handlers using open cab airblast methods in pome fruit multiplied by two uncertainty factors of 10 is also compared*.

Pesticide	NOAEL (mg/kg/day)	10% BMD	10% BMDL	1 SD-based BMD	1 SD-based BMDL	20% BMD	20% BMDL	Hybrid BMD	Hybrid BMDL	Normalized Effect Size (%)	NES BMD	NES BMDL	Maximum endpoint value
Acetamiprid	10	1.99	0.30	44.62	26.74	--	--	1.18	0.16	30.86	12.09	2.23	8.6
Azinphos methyl	0.15	0.23	0.17	4.42	3.51	0.50	0.35	0.75	0.4	25.99	0.633	0.469	0.09
Emamectin benzoate	0.075	0.91	0.19	--	--	--	--	--	--	--	--	--	--
Methoxyfenozide	16.8	27.69	9.56	28.93	10.13	--	--	83.11	10.16	20.75	36.21	5.84	4.52
Novaluron	4.38	2.00	0.90	7.66	2.92	--	--	21.28	2.34	25.44	6.24	2.75	6.13
Phosmet	4.5	2.75	0.58	4.37	3.97	4.60	4.53	7.36	5.89	25.99	4.68	4.62	188.20
Spinetoram	2.7	6.62	3.31	--	--	--	--	--	--	--	--	--	--
Thiacloprid	1.2	1.40	0.64	--	--	--	--	--	--	--	--	--	--

*The Maximum endpoint value represents the minimum point of departure in mg/kg/day which would, when combined with the deterministic dose value used in the latest EPA exposure assessment, permit regulatory approval of the compound based on occupational risks.

Table 3:

Ratios of points of departure and points of departure vs $dose_{EPA}$, i.e., the EPA-derived deterministic doses used in the human health risk assessment for pesticide handlers. Two values are presented for phosmet: oral is calculated using the oral dosing study values compared with doses calculated using an adjustment factor for dermal absorption. The dermal study does not include adjustment for dermal absorption and uses the dermal toxicity testing data.

	BMD	***BMD***	***BMDL***	***BMD***	***BMDL***	***NOAEL***	**BMD_{NES}**	**BMD_{NES}**
	BMDL	***NOAEL***	***NOAEL***	***$Dose_{EPA}$***	***$Dose_{EPA}$***	***$Dose_{EPA}$***	**$BMDL_{NES}$**	***NOAEL***
Acetamiprid	5.76	0.17	0.03	226	39	1302	7.44	1.21
Azinphos methyl	1.40	3.32	2.36	60	43	18	1.35	4.22
Emamectin benzoate	4.83	12.15	2.52	10171	2106	837		
Methoxyfenozide	2.90	1.65	0.57	8073	2787	4898	2.99	0.84
Novaluron	2.22	0.46	0.21	249	112	545	2.27	1.42
Phosmet								
Oral	1.02	1.53	1.01	153	151	92	1.01	1.04
Dermal	3.86	0.90	0.23	175	45	270	2.19	1.39
Spinetoram	2.00	2.45	1.23	23629	11820	9643		
Thiacloprid	2.20	1.17	0.53	368	167	314		

Table 4:

Exceedance fractions of probabilistically-estimated doses for each compound and the associated points of departure estimated from the dose-response studies, and the ratio of these exceedance fractions, demonstrating the difference in protective ability of each selected point of departure as a regulatory limit

	Exceedance fractions			Ratio of Exceedance Fractions		Margin of Exposure: mean (% over LOC)			
	NOAEL	BMD	BMDL	BMDL/BMD	BMD/NOAEL	NOAEL	BMD	BMDL	SD-BMD
Acetamiprid	5.0	24.8	61.2	2.5	5.0	1502983 (5)	28352 (25)	343 (62)	51011 (0.7)
Azinphos methyl	72.2	55.7	63.5	1.1	0.8	432 (54)	446 (44)	330 (60)	8571 (5)
Emamectin benzoate	19.6	2.0	9.7	4.8	0.1	53581 (20)	650122 (2)	135740 (10)	-
Methoxyfenozide	0.005	0.002	0.015	8.4	0.4	16573165 (0.002)	27316127 (0.0006)	9430920 (0.006)	28539384 (0.0005)
Methoxyfenozide (dermal + inhalation routes)	2.7	1.5	4.7	3.1	0.6	19248 (0.6)	31725 (0.25)	10953 (2)	33146 (0.23)
Novaluron	5.09	9.3	15.7	1.7	1.8	36101 (1)	16485 (3)	7418 (8)	63136 (0.5)
Phosmet (oral)	8.94	13.5	36.9	2.7	1.5	12945 (4)	111025 (15)	1668 (26)	12571 (4)
Phosmet (dermal)	12.62	13.8	34.9	2.5	1.1	13620 (2)	8323 (3)	1756 (20)	13227 (1.7)
Spinetoram	0.6	0.2	0.4	2.7	0.3	8542041 (0.53)	20943820 (0.15)	10471910 (0.4)	-
Spinetoram (dermal + inhalation routes)	22.6	15.3	20.8	1.4	0.7	45282 (22)	1283 (4)	55512 (20)	-
Thiacloprid	32.0	30.1	40.2	1.3	0.9	138 (88)	161 (87)	74 (94)	-

*Exceedance fractions are the proportion of the dose distribution above the deterministic point of comparison of the NOAEL, BMD, or BMDL. In the MOE, exceedance fraction is the proportion of the distribution below the LOC (100 for all cases but Emamectin Benzoate, which was 300).