ACROPOLIS

Aggregate and Cumulative Risk Of Pesticides: an On-Line
Integrated Strategy
SEVENTH FRAMEWORK PROGRAMME

Deliverable 5.5   A scientific paper describing the integrated risk model and its statistical validation

Biometris, Wageningen University and Research Centre (WUR), Wageningen, The Netherlands

Contact person: Hilko van der Voet, hilko.vandervoet@wur.nl

Project team Acropolis: Hilko van der Voet (Biometris/WUR), Waldo J. de Boer (Biometris/WUR), Johannes Kruisselbrink (Biometris/WUR), Paul W. Goedhart (Biometris/WUR), Gerie W.A.M. van der Heijden (Biometris/WUR), Marc C. Kennedy (Fera), Polly E. Boon (RIVM), Jacob D. van Klaveren (RIVM)

Table of Contents

The MCRA model for probabilistic single-compound and cumulative risk assessment of pesticides

Hilko van der Voet [a, *], Waldo J. de Boer [a], Johannes Kruisselbrink [a], Paul W. Goedhart [a], Gerie W.A.M. van der Heijden [a], Marc C. Kennedy [b], Polly E. Boon [c], Jacob D. van Klaveren [c]

[a] *Biometris, Wageningen University and Research Centre (WUR), Wageningen, The Netherlands*

[b] *Food and Environmental Research Agency (FERA), Sand Hutton, United Kingdom*

[c] *National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands*

 * author for correspondence

Abstract

Pesticide risk assessment is hampered by worst-case assumptions leading to overly pessimistic assessments. On the other hand, the cumulative health effects of similar pesticides are often not taken into account because of the lack of suitable models for practical applications. This paper describes models and a web-based software system developed in the European research project ACROPOLIS. The models are appropriate for both acute and chronic exposure assessments of single compounds and of multiple compounds in cumulative assessment groups. The software system called MCRA (for Monte Carlo Risk Assessment) is available for stakeholders in pesticide risk assessment. We describe the MCRA implementation of the methods as advised in the 2012 EFSA Guidance on probabilistic modelling, as well as more refined methods which were developed in the ACROPOLIS project. In this paper the various models are described, with an emphasis on methods for cumulative assessments. Two main approaches, sample-based or compound-based, are contrasted. It is shown that additional data on agricultural use of pesticides may be helpful in obtaining more realistic risk assessments. Examples are given of the model and software validation, using both simulated data and comparisons against the standard software DEEM-FCID used by the Environmental Protection Agency in the USA. It is shown that the EFSA Guidance pessimistic model may not give an appropriate modelling of exposure.

1   Introduction

Cumulative effects on health from multiple similar pesticides ('cocktail' effects) have been overlooked in traditional risk assessments, which only considered single compounds. The European Parliament and the European Council have regulated that for risk assessments leading to decisions on applications concerning maximum residue limits (MRLs) 'account shall be taken of the possible presence of pesticide residues arising from sources other than current plant protection uses of active substances, and their known cumulative and synergistic effects, when the methods to assess such effects are available' (EC, 2005). In response to this perceived lack of sufficient methodology for cumulative risk assessment (Boobis et al. 2008) the EU-funded project ACROPOLIS (www.acropolis-eu.com) has developed a web-based tool for probabilistic cumulative exposure assessment, called MCRA (for Monte Carlo Risk Assessment, available at mcra.rivm.nl). The software also facilitates aggregate exposure assessments, on which as reported in a companion paper (Kennedy et al., in prep.).

Different active substances can have the same health effect, and this is commonly expected for substances within a chemical class. For example, many pesticides in the triazole group can lead to neurological damage in embryos. Therefore the European Food Safety Authority (EFSA) has defined Cumulative Assessment Groups (CAGs) containing 7 triazole compounds for acute and 11 triazole compounds for chronic risk assessment (EFSA 2008). Dose-effect studies have been performed to estimate Relative Potency Factors (RPFs), where flusilazole and cyproconazole were the reference compounds for acute and chronic, respectively.

Cumulative effects were also considered in the EFSA guidance on probabilistic modelling (EFSA 2012). In this work a tiered approach was proposed where basic probabilistic assessments would comprise two alternative model runs. Pessimistic and optimistic model runs treat major uncertainties using assumptions expected to give over-estimation and under-estimation of exposure, respectively. The guidance document describes at length how the basic model runs should be performed, both for single-compound and cumulative assessments. In a higher tier refined probabilistic assessments might be needed based on the results of the two basic model runs and additional sensitivity analyses . In this paper the implementation of the EFSA guidance in the MCRA model is described. The application to triazole exposure assessment in eight European countries is reported in a companion paper (Boon et al., in prep.).

Previous approaches to cumulative risk modelling have been described. Price and Chaisson (2005) presented a conceptual framework for modelling aggregate and cumulative exposures, stressing the importance of person-oriented modelling. Meek et al. (2011) described a WHO/IPCS framework with multiple tiers for both exposure and hazard modelling, where tier 3 corresponds to probabilistic approaches. In practical work, probabilistic cumulative exposure assessments have usually been implemented as index compound equivalents assessments on RPF-weighted total exposure per sample (EPA 2002, Boon and van Klaveren 2003, Caldas et al. 2006ab, Boon et al. 2008, Jensen et al. 2013). Van Klaveren et al. (2009) compared this sample-based approach to an alternative approach based on modelling each compound separately followed by a summation of the single-compound exposure distributions.

This paper describes the statistical models in the web-based MCRA system for single-compound and cumulative risk assessment, shows some example calculations for the triazole group, and validation results.

2   Methods

MCRA 8.0 is a research tool implementing many possibilities for modelling single-compound or cumulative exposure, including the procedures as described in the EFSA guidance for acute or chronic exposure assessment according to both an optimistic and a pessimistic model run.

One-day (24 hour) data on food consumption and per-sample residue concentration data are taken as inputs? to characterise the short-term exposure distribution relevant for acute risks. The consumption and concentration distributions are integrated by Monte Carlo (MC) sampling. The number of MC iterations can be chosen, a number of 100,000 is typically found to be sufficient.

Estimates of the long-term exposure distribution relevant for chronic risks are obtained from statistical modelling of the 24-hour consumptions multiplied by the average residue concentration, without further modelling. Application of a random or mixed effects model to transformed data separates within- and between-individual variation in exposure. Pesticide exposures per day are often 0, which necessitates the use of two-part models for the frequency of exposure and the amount of exposure, respectively (Dodd et al. 2006, de Boer et al. 2009).

Uncertainty analysis can provide uncertainty limits on exposure percentiles and other statistics in a two-dimensional Monte Carlo approach where datasets are bootstrapped in an outer loop. Uncertainty aspects of MCRA are more fully discussed in a companion paper (Kennedy et al., in prep.).

2.1   Food code conversions and processing factors

Food codes used in food consumption surveys are typically not the same as food codes used in monitoring programs. For example, the consumption survey may record 'apple pie', whereas pesticide measurements have been made on the raw agricultural commodity (RAC) 'apple'. This requires at least two conversion steps: a recipe of apple pie should specify the proportion of 'peeled apple', and subsequently a processing factor may be known for the residue concentration in 'peeled apple' relative to the concentration in the (unpeeled) RAC 'apple'. MCRA allows the input of both the recipe proportions and the processing factors in relation to a translation of food codes, thus linking foods-as-eaten to foods-as-measured.

Processing factors are defined as the concentration in the processed food divided by the concentration in the unprocessed food. A database of processing factors is available at
http://www.bfr.bund.de/cm/349/bfr-compilation-of-processing-factors-for-pesticide-residues.zip or
http://chemkap.rivm.nl/groente-fruit/processing-factors/.

Concentration changes can be due to 1) chemical degradation, e.g. in heat treatments, and/or 2) weight changes of the food, e.g. drying of grapes to raisins. Care has to be taken that weight changes are not accounted for in both the recipe and the processing factor data, and therefore would be double-counted. If the same food code translation is made in a recipe and in the table of processing factors then MCRA will adjust the processing factor for the weight change such that only the chemical degradation is reflected in the processing factors that will be applied to the residue concentration data of the RAC.

Processing effects are usually considered as fixed values, but in principle they can be variable and/or uncertain. In MCRA both the variability and the uncertainty of processing factors can be modelled using logistic-normal (for processing factors restricted to the interval [0,1]) or log-normal distributions (see appendix in van der Voet et al. 2009 for details).

2.2   Modelling residue concentrations per food

For modelling of acute risk it is necessary to model distributions of residue concentrations from which values can be sampled in the Monte Carlo simulations. For modelling of chronic risk it is sufficient to work with mean residue concentrations, but for non-trivial data, e.g. where not all measurements are positive values, these are still dependent on the type of model assumed.

The two basic concentration models in MCRA are the empirical model and the lognormal model. The empirical model (EFSA optimistic) uses the data as they are, whereas the lognormal model (EFSA pessimistic) fits a lognormal distribution to the available positive residues.

Typically, residue concentration datasets contain only a limited number of positive values and a large number of 'non-detects' or 'less-thans', i.e. measurements censored at a limit value, which depending on the reporting institution may be a formal limit of detection, limit of quantification, or other value. For the purpose of exposure modelling any such limit will be referred to as the Limit Of Reporting (LOR). We will refer to values' <LOR' as non-detects (NDs). Note, that LOR may be in principle be different for each measurement, but in practice it is often a constant value for a specific analytical method as applied in a specific laboratory and a specific time period.

For both the empirical and the lognormal model a decision is needed how to handle the NDs. MCRA implements simple imputation options, such as replacing NDs by zero (EFSA optimistic) or replacing NDs by the LOR (EFSA pessimistic). The EFSA pessimistic model is therefore a mixture distribution of a spike of NDs imputed at LOR and a lognormal distribution fitted to the positive values (lognormal ND-spike model). In more complex models it is also possible to consider the NDs as a censored tail of the lognormal distribution and fit a censored lognormal distribution (censored lognormal model). An even more complex model is to assume that the data are a mixture of a certain percentage of true zero concentrations and positive values which may partly be censored. If there are sufficient positive values, the assumption of lognormality may then be used to estimate the percentage of true zeroes as well as the parameters of the lognormal distribution (censored lognormal zero-spike model).

Sometimes more knowledge is available from other sources. When agricultural use of a specific compound on the commodity is known to be unauthorised, then NDs will be set to 0 mg/kg (note that positive values in this case would represent unauthorised use). When agricultural use is authorised, then it may be the case that a value for the percentage crop treated (%CT) is available. In that case, an appropriate concentration model is that (100-%CT)% of the concentrations is actually 0, and that NDs are a mixture of true zeroes and censored positive values. If the actual percentage of positive residues in the dataset exceeds the %CT value, the latter will be upwardly adjusted. In summary, depending on the information of agricultural use that is available, a certain fraction of true zero concentrations is assumed, and the remaining fraction of the data is modelled as sketched above, including the ND imputation options.

Agricultural use information (authorization status and %CT) may be available at a general level or country- and/or time-specific. In the latter case, MCRA allows to weigh the %CT data from different locations in proportion to the sampling locations indicated in the residue concentration dataset.

Finally, another type of imputation may be needed in cumulative exposure assessments, where not all compounds have been measured on each sample because of different analytical methods, and therefore the data contain missing values (MVs). As described further in section 2.4, MVs are imputed by making a draw from the appropriate distribution as found in the above modelling. This distribution is therefore in principle a mixture of true zeroes and the fitted distribution for positives, either empirical or lognormal.

The approach to uncertainty analysis depends on the chosen concentration model. In most cases a non-parametric bootstrap is used to generate new consumption and concentration datasets. In the case of the lognormal ND-spike model it is also feasible to specify a parametric uncertainty model as is also described in the EFSA Guidance, using a beta distribution for the uncertainty in the fraction of positives, an inverse chi-square distribution for the uncertainty in the variance and a normal distribution for the uncertainty in the mean on the log scale.

2.3   Modelling single-compound exposure

For acute risk assessment a Monte Carlo procedure combines random draws of consumption patterns (individual-days) with random values sampled from the residue concentration distribution for each relevant food. The residue concentrations are multiplied by a processing factor and/or a stochastic variability factor if needed. Processing factors were already discussed in section 2.1. Stochastic variability factors are the probabilistic equivalent of the unit variability factors which have been introduced in pesticide exposure modelling to account for the fact that RAC measurements are

normally made on composite samples (say 12 apples mixed together), whereas consumers are confronted with single units (e.g. one apple). Residue concentrations in composite samples under-estimate the variation. Variability factors are defined as the 97.5[th] percentile divided by the mean of the distribution of residue found in single units. Stochastic variability factors are random draws from a suitable distribution characterised by a mean that is equal to a sampled composite residue, and a given ratio p97.5/mean. In MCRA the lognormal distribution or a scaled beta distribution can be chosen. The motivation for the latter is that a single unit in a composite sample cannot contain more residue than the composite sample as a whole, and therefore there is a maximum to the possible concentration. For example, if all residue in a composite sample of 12 apples would reside in one 'hot' apple, then the concentration would be 12 times the measured composite sample concentration.

Exposures y_i for an individual-day i are simulated according

$$y_i = \sum_{j}^{foods} \sum_{h}^{proctypes} \sum_{u}^{units} x_{ijhu} \cdot svf_{ijhu} \cdot pf_{jh} \cdot c_{ijh} \bigg/ BW_i$$

where the consumption x_{ijhu} relates to unit portion u of food j of processing type h. Individual consumed amounts x_{ijhu} are derived from resampled consumption patterns x_{ijh} and a dataset of unit weights per food type. Resampling can use sampling weights w_i when available in the food consumption data. The individual has body weight BW_i. Residue concentrations c_{ijh} are sampled from the residue concentration distribution fitted before, and multiplied by the processing factor pf_{jh}. Note that separate draws from the same distribution are made for different processing types of the same food, but that different units u in the same consumption dataset entry start from the same sampled residue value, which is then modified by a stochastic variability factor svf_{ijhu} sampled from the unit variability distribution.

For chronic risk assessment, different models to assess the long-term exposure (also called usual or habitual intake) may be employed. It is well-known that direct use of short-term (e.g. per-day) data leads to an over-estimation of upper tails in chronic risk assessment (e.g. Dodd et al. 2006). MCRA implements various models for this, such as the BetaBinomial-Normal (BBN) model and the Iowa State University Foods (ISUF) model, as were previously described (de Boer et al. 2009). More recently, the Logistic-Normal-Normal (LNN) model was added, which is very similar to the National Cancer Institute (NCI) model (Tooze et al. 2006, Goedhart et al. 2012). However, all these models are based on assumptions about the statistical distribution of exposures, which never can be fully tested in practice. For this reason, and to avoid unwanted over-optimism, the EFSA Guidance prescribed the use of a simpler non-parametric method for chronic assessments. In this Observed Individual Means (OIM) method, the intakes calculated for the different days of a person are just averaged to obtain an estimated long-term exposure distribution. This has been shown to give very pessimistic estimates of high tail exposures (Goedhart et al. 2012).

2.4   Modelling cumulative exposure

In cumulative acute risk assessment two main approaches have been identified (van Klaveren et al. 2009), which are both available in MCRA. In Approach 1, or the **sample-based approach**, residue concentrations are recalculated in terms of an index compound, and then summed:

$$y_i = \sum_{j}^{foods} \sum_{h}^{proctypes} \sum_{u}^{units} x_{ijhu} \cdot \sum_{k}^{compounds} \left(RPF_k \cdot pf_{jhk} \cdot svf_{ijhuk} \cdot c_{ijhuk} \right) \bigg/ BW_i$$

In Approach 2, or the **compound-based approach**, exposures are calculated independently for each compound, and then the cumulative exposure is calculated as a weighted sum:

$$y_i = \sum_{k}^{compounds} \left(RPF_k \cdot \sum_{j}^{foods} \sum_{h}^{proctypes} \sum_{u}^{units} x_{ijhu} \cdot pf_{jhk} \cdot svf_{ijhuk} \cdot c_{ijhuk} \right) \Big/ BW_i$$

In both equations k identifies the compound, and RPF_k is the Relative Potency Factor for compound k relative to the index compound.

In Approach 1 correlations between compounds in the residue concentrations are retained simply by resampling concentration vectors for all compounds simultaneously from the set of samples. Residue concentrations are made comparable between compounds by weighing with relative potency factors (RPFs), and may be modified by applying processing factors pf and/or stochastic variability factors svf.

A simplified version of the sample-based approach uses the same processing factors pf and stochastic variability factors svf for all compounds k. Then the sample-based approach can be performed whenever a program for single-compound probabilistic exposure assessment is available. It is only necessary to calculate the RPF-weighted sums of concentrations in the chemical samples as input for a standard single-compound assessment. This simplified approach has been used previously in several studies (Caldas et al. 2006, Boon et al. 2008).

In some datasets not all compounds of the CAG are measured in all samples. Approach 2 avoids the assumption of zero concentrations in non-measured samples needed in Approach 1. Approach 2 can also be used when the residue data are only available without sample identification. Approach 2 was used in van Klaveren et al. (2009) because the assumption of zero values for unmeasured residues was shown to lead too under-estimation.

The EFSA Guidance (EFSA 2012) has proposed the use of the sample-based approach for basic cumulative acute assessments, using an imputation method to replace missing values in the data matrix for each food. For basic optimistic model runs missing values are replaced by zero. For basic pessimistic model runs, after setting all residue levels below the LOR to LOR, provisional RPF-weighted sums per sample are used for sorting the samples from highest to lowest cumulative exposure. For each pesticide-RAC combination as many residues as there are MVs are sampled from the already fitted binomial-lognormal distribution, and also these are sorted from high to low. Then the first MV is imputed by the highest sampled residue, and the RPF-weighted sum and sorting of samples is adapted if necessary. This is continued until all sampled residues have been used and all MVs have been imputed.

When data on agricultural uses (AUs) (authorization status and/or %CT) are available, the imputation procedure has still to be adapted. The EFSA Guidance only uses available knowledge about the authorization status of agricultural uses (combinations of pesticides) to prevent unauthorised combinations to be generated in the imputation (a zero value is imputed instead). In the Acropolis project a more complete model was developed. With p compounds in a CAG there are in principle 2^p-1 possible AUs (combinations), with between 1 and p pesticides. In practice, only a limited number (n_{AAU}) of AUs is authorised, and the remaining $n_{NAAU} = 2^p - n_{AAU} - 1$ non-authorised AUs can be lumped into one group. For example, if a CAG consists of 3 compounds (A, B, C) and the AU data specify 10% CT for A and 5% CT for A+B, then the non-authorised AUs are B alone, C alone, A+C, B+C and A+B+C. In $100 - 10 - 5 = 85\%$ of cases the crop can be considered to be untreated with any compound if authorisation status is used. On this assumption we can also derive %CT per compound (%CT$_{compound}$), which in the example will be $10 + 5 = 15\%$ for compound A, 5% for compound B and 0% for compound C.

If AU authorization data is available, then there are still four ways to obtain residue values in the modelling, depending on the dichotomies sample- or compound-based and use %CT data or not. In a sample-based acute cumulative assessment the imputation of MVs will be a value from the residue model with probability %CT$_{compound}$ and zero otherwise. If no %CT data is used, then there are only zeroes for the non-authorised combinations (EFSA Guidance method). In a compound-based acute

cumulative assessment Monte Carlo simulations start by sampling an AU for each simulated sample. This already fixes zero residue values for the compounds which are not in the sampled AU. For the compounds included in the AU a value from the residue model is simulated with probability %CT$_{compound}$ (100% if no %CT data are used), and zero otherwise.

Whenever pesticides are used together, their levels may be correlated. This is automatically addressed in the sample-based approach, but not in the compound-based approach. However, co-occurrences of positive values in practical datasets are often so sparse, that correlations between residues are difficult to estimate. As an alternative, it is possible to perform in MCRA a sensitivity analysis. In the standard analysis residues are sampled independently (zero correlation). In an alternative run the sampled residues within each sampled AU group are re-ordered according to their ranks. So in the A+B group the highest sampled residue for A will coincide with the highest sampled residue for B. This should give a more pessimistic estimate in the upper tail if the food under consideration is a risk driver.

In cumulative chronic risk assessment the same two approaches (sample or compound based) can be followed, now using the mean residue concentrations (derived from the chosen concentration models) instead of sampled concentration values. The issues are partly simpler, e.g. unit variability and correlations between residues have not to be addressed.

2.5 Validation of models and software

Validation is the process of evaluating models and software to determine whether they are fit for purpose. In probabilistic exposure assessment of pesticides the main purpose is to obtain reliable estimates of the upper tail percentiles of the exposure distribution. In addition, the purpose is to estimate reliably the contributions to exposure from different sources (foods in the diet, and, in aggregate assessments, also dermal and inhalatory contributions) and, in cumulative assessments, from different compounds.

Validation compares model output to reference values. The reference values may be true values or values from a historically accepted procedure. A practical validation of dietary exposure assessments models against true values can be obtained from a duplicate diet study (Boon et al. 2003). Given the high costs and low precision of such a study it is often preferable to obtain true values (and model input data) from a simulation model. Historically accepted procedures for exposure assessment are typically other software tools which have been used for a longer time than the software to be evaluated.

In the current study MCRA 8.0 was validated against true simulated values at several levels of complexity, and against results from the previous release of the same software (MCRA 7.1) and the DEEM-FCID 3.16 software used by the Environmental Protection Agency in the United States (EPA-US).

2.5.1 Synthetic data validation

The computer code of MCRA 8.0 includes a large number of automatic tests that run on a set of prepared synthetic datasets which are simple enough to know the desired outcome. These tests are ordered in three categories: 1. Data integrity tests to validate that all input is correctly transferred to the internal database structures, 2. Unit and modular tests to validate the outcomes of specific calculations, and 3. Scenario tests to validate the outcomes of complete runs of the program. A test report, generated from a run of all tests, is generated automatically, and is available at any time.

As an example we describe here a scenario test written to validate a cumulative exposure assessment where additional data on agricultural use are available. This test can also be used to illustrate the potential advantage of the models described in section 2.4. The consumption data for this test are simplified to just two individuals (with body weights 65 and 75 kg) who both consume 100 g of banana on each of the two survey days. The concentration data describe 100 samples all measured with an analytical method which can measure two compounds, bitertanol and triadimefon, with an

LOR of 0.05 mg/kg for the non-detects. In one sample both compounds are found (bitertanol at 0.08 mg/kg, triadimefon at 0.03 mg/kg, note that individual positive values can be below the LOR for non-detects). In four more samples only bitertanol is found (at levels 0.12, 0.10, 0.10 and 0.10 mg/kg) and in one other sample only triadimefon is found (at 0.07 mg/kg). The agricultural use data specify three authorised uses: a formulation with only bitertanol is used in 10% of the crop, a formulation with only triadimefon in 5% of the crop, and a formulation containing both compounds in 2% of the crop. The RPFs of bitertanol and triadimefon are 2.1 and 1.2, respective to the index compound flusilazole.

Given these input data it is possible to calculate by hand the expected 95% percentile of exposure under various models, and these were compared with the results from the program.

2.5.2   Validation against other software

MCRA 8.0 was validated against results from the previous release of the same software (MCRA 7.1) and against the DEEM-FCID 3.16 software used by the Environmental Protection Agency in the United States (EPA-US). As an example the latter validation is reported here. It is in the interest of future harmonization to compare the results obtained with US and European software tools using the same basic data. DEEM-FCID is the standard program for the exposure assessment of pesticides in the US. Models for exposure assessment of pesticides have developed quite independently on both sides of the Atlantic. For example, the US DEEM model incorporates a fixed US consumption database, whereas the EU MCRA model allows users to link any consumption database. Obviously then, given the architecture of DEEM the validation has to use the US WWEIA-FCID consumption data.

The DEEM-FCID Version 3.16 software was downloaded from http://www.epa.gov/pesticides/science/deem/. This version incorporates the NHANES What We Eat in America – Food Commodity Intake Database (WWEIA-FCID) dietary survey data for the years 2003-2008. The consumption data are separately available at http://fcid.foodrisk.org/.

For the calculations with MCRA the WWEIA-FCID 2003-08 food consumption WWEIA codes were translated to FCID codes using the Recipes table. Several databases were constructed which differ in the residue concentrations. As a first check an artificial value 1 was used for nine FCID codes of apple products to generate a consumption distribution for the individual-days in the WWEIA dataset including sampling weights. Secondly, a true Monte Carlo analysis was performed to calculate the captan exposure distribution using Dutch captan monitoring data for apple (n=260). For this the nine FCID codes for apple products were linked to the single Dutch Raw Agricultural Commodity (RAC) code for apple. Processing factors were not available for the FCID codes within the Acropolis project, therefore validation was restricted to the basic exposure calculation.

3   Results

3.1.1   Synthetic data validation cumulative exposure

The first step in cumulative exposure modelling is the modelling of the cumulative residue concentrations expressed as equivalents of the index compound flusilazole. The EFSA basic pessimistic model requires to replace the non-detects by the LOR, and fit the weighted sum by a lognormal distribution. The results in Figure 1a show that this may lead to a high spike in the flusilazole equivalents distribution corresponding to the samples where only non-detects have been registered. The lognormal distribution could be reasonable for the measured compounds, but gives an inappropriate fit to the flusilazole equivalents distribution, and this leads to unconservative estimates of the upper tail of the distribution. Data on the percentages crop treated (agricultural use data) can at least partly remedy this problem as shown in Figure 1b. In accordance with the supplied agricultural use data 88% of samples is free of bitertanol and 93% is free of triadimefon. Recognizing that a large part of the samples is untreated (and thus corresponds to true zero values) leads in this case to 83% zero concentrations and a more acceptable fit of the lognormal model.

Using the flusilazole equivalents distribution the cumulative exposure has been estimated. The EFSA basic optimistic model replaces non-detects by zero, and therefore the flusilazole equivalents distribution consists of a spike of 94 % zeroes plus a pattern of spikes which exactly corresponds with the values predicted by hand calculation from all combinations of concentrations on samples and body weights in the input data (Figure 2a). The cumulative exposure estimated with EFSA basic pessimistic model is again dependent on the use of agricultural use data. Without such data all estimated exposures are positive because in this example banana was consumed on all days, and the flusilazole equivalents distribution also had no zeroes. Based on the agricultural use data there are only 17% positives, and the estimated distribution is broader. The estimated 99.9th percentiles for the three cases were 0.39 (optimistic), 0.35 (pessimistic, no AU) and 0.58 (pessimistic, AU) µg/kgBW/d (Table 3).
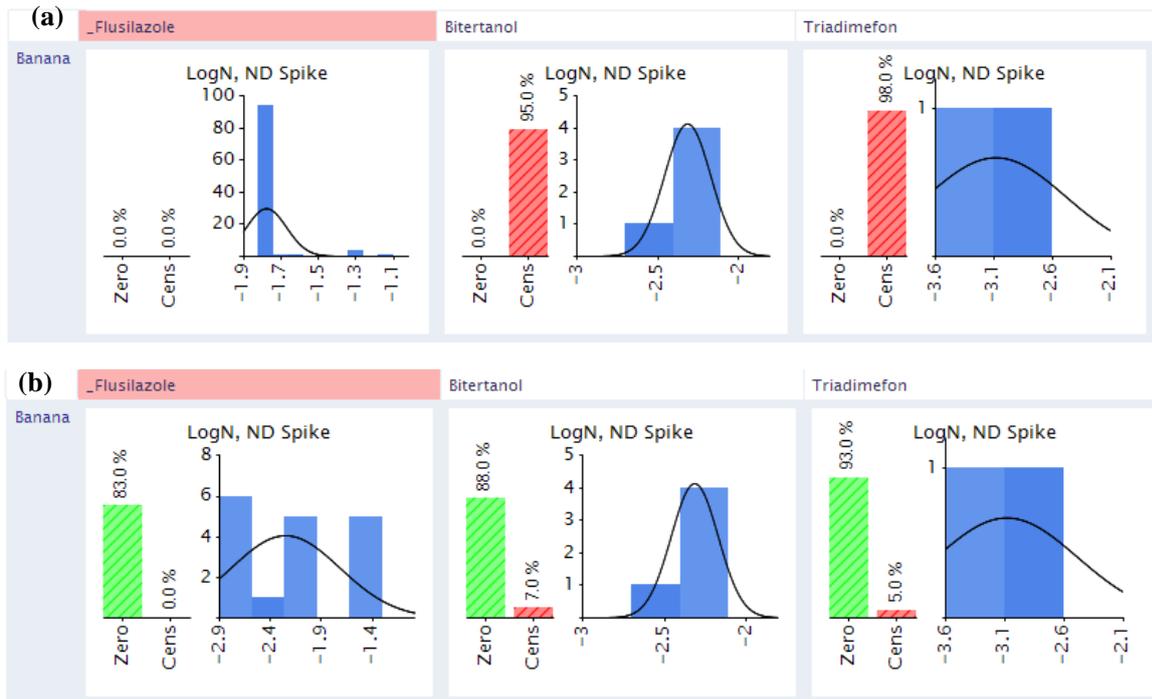

Figure 1. Synthetic example. MCRA output showing histograms of ln(concentration) of calculated flusilazole equivalents (left) from measured bitertanol (middle) and triadimefon (right) in banana. Calculation according to EFSA Guidance, basic pessimistic model (fit of LogNormal Non-Detect Spike model is shown), (a) without and (b) with using data on agricultural use.
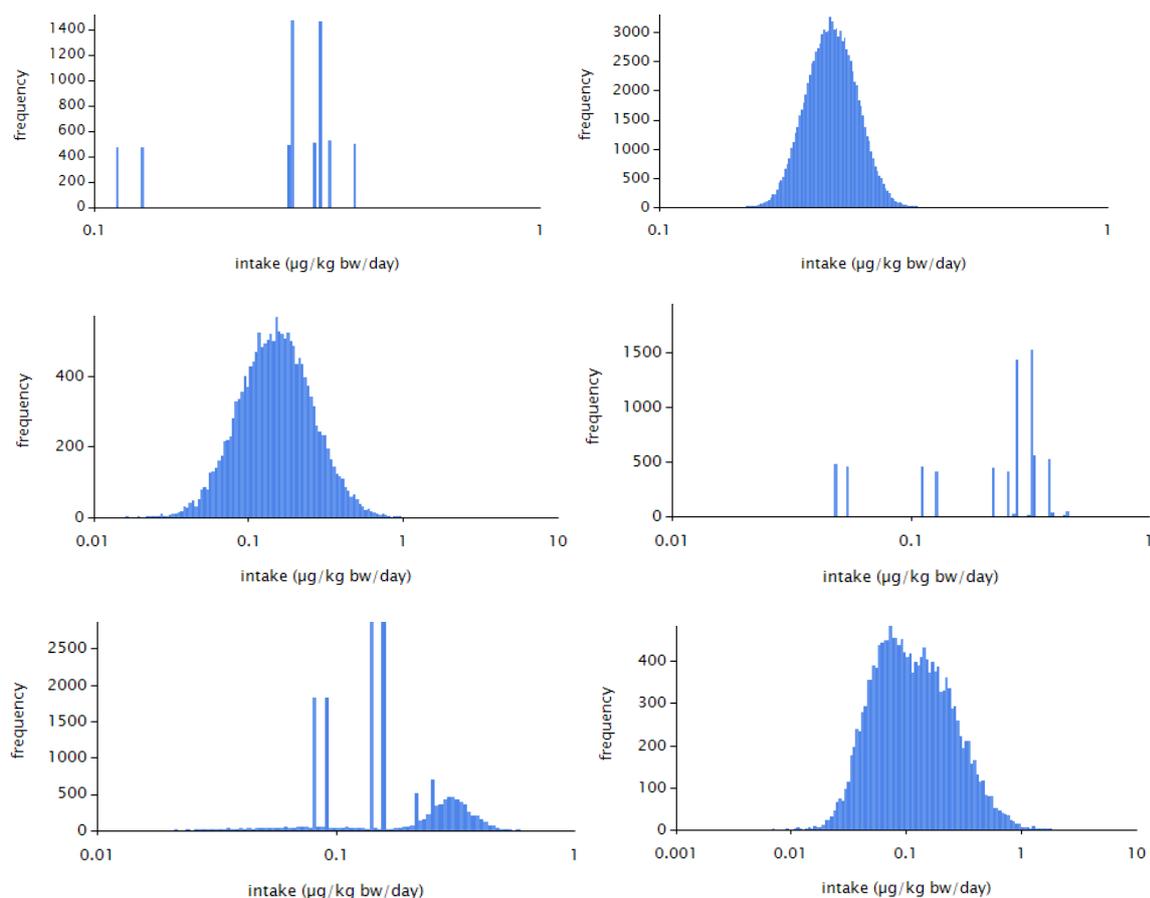
Figure 2. Synthetic example. MCRA output showing simulated positive exposures in 100,000 Monte Carlo runs according to (a) EFSA Guidance optimistic, 6% positive; (b) EFSA Guidance pessimistic, AU data not used, 100% positive; (c) EFSA Guidance pessimistic, AU data used, 17% positive; (d) compound-based empirical, NDs set to 0, 7% positive; (e) compound-based ND spike lognormal, NDs set to LOR, AU data used, 17% positive; (f) compound-based censored lognormal, AU data used, 17% positive.

Table 1. Synthetic example. Summary statistics for exposure distributions using various probabilistic methods. ND = non-detect, LOR = limit of reporting, AU = agricultural use, maxcor = maximized correlation in sensitivity analysis (see text).

	Positive exposure (% of population)	99.9th percentile (µg/kgBW/d)	Exceedance of 0.5 µg/kgBW/d (% of population)
Sample based cumulation			
EFSA optimistic	6	0.39	0
EFSA pessimistic	100	0.35	0
EFSA pessimistic + AU	17	0.58	0.2
Compound-based cumulation			
Empirical, ND→0	7	0.39	0
Empirical, ND→0, maxcor	7	0.39	0
NDSpike-LogN, ND→LOR + AU	17	0.46	0.03
NDSpike-LogN, ND→LOR, maxcor + AU	17	0.50	0.09
CensLogN, + AU	17	0.81	0.5
CensLogN, maxcor + AU	17	0.84	0.6

In Figure 2 and Table 1 also some results for alternative cumulative methods are shown, which are variations of Approach 2 (compound based). Using empirical distributions and setting non-detects to zero (a compound-based 'optimistic' approach) gives very similar results as the EFSA sample-based optimistic approach: there are 7% positive exposures as a result of randomly combining 5% positive

bitertanol with 2% positive triadimefon exposures (1-0.95x0.98). The distribution may look different because other combinations can be made, e.g. there are now results below 0.1 µg/kgBW/d because the lowest triadimefon residue could be sampled without the positive bitertanol value which was actually found on the same sample. In a compound-based model where AU data are combined with the ND-spike LogNormal model, and NDs are imputed by the LOR (this is the same concentration model as in the EFSA pessimistic model) a slightly lower tail percentile is found than with the AU-adapted EFSA pessimistic model. The exposure distribution however looks very weird because of the spikes caused by the imputed non-detects (10% of the population). An alternative censored lognormal model combined with AU data in the same way results in a more reasonably looking exposure distribution, and a higher tail percentile (0.81 µg/kg bw/d).

A sensitivity analysis was performed to correct for the lack of modelling correlation between pesticides in the compound-based approach. The results, shown in Table 1, show only slightly higher tail percentiles when the sampled residues are paired based on their ranking.

3.1.2 Validation MCRA against DEEM

The percentiles calculated with MCRA 8.0 and DEEM 3.16 for the total consumption distribution of apple products are reported in Table 2, and differed by less than 0.5%, except for the very low exposure at p70 (6% difference) and in the extreme upper tail, at p99.9 (0.6% difference).

The differences between MCRA and DEEM may be related to technical differences in the algorithms. Whereas MCRA stores all the exposures for the simulated individual-days, the older DEEM program applies a binning algorithm to reduce storage requirements and gain speed. In the interest of precision for risk assessments the bins are finer in the upper tail and become wider at lower exposures. The consequence of this is seen in Figure 3: MCRA shows a trimodal distribution for the 42.6 % positive consumptions, but in DEEM the two lower components are not distinguished.

Table 2. Consumption US population of apple products (g/kgBW/day). Percentiles MCRA and DEEM.

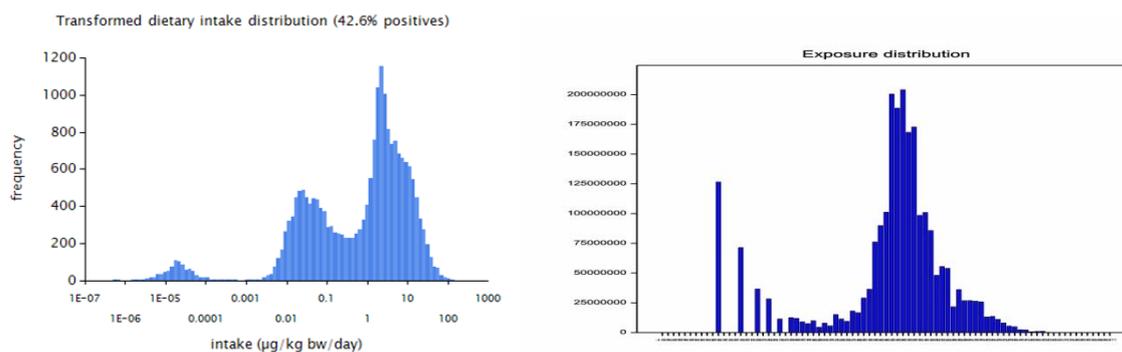| Percentage	MCRA	DEEM
60.00 %	0	0
70.00 %	0.02788	0.02961
80.00 %	0.7995	0.8027
90.00 %	2.863	2.871
95.00 %	6.079	6.075
97.50 %	11.17	11.17
99.00 %	19.89	19.85
99.50 %	27.84	27.83
99.75 %	37.85	37.92
99.90 %	53.79	54.09

Figure 3. WWEIA-FCID data for nine apple products. Total consumption distribution according to (a) MCRA and (b) DEEM.

In the second validation a full two-dimensional Monte Carlo assessment was run with MCRA. Individual-days and captan concentrations were resampled in the inner loop of 100,000 Monte Carlo iterations and percentiles were calculated. In the outer loop the datasets themselves were bootstrapped 100 times to generate uncertainty limits on the percentiles. The comparison with DEEM (Table 3) showed DEEM estimates to fall between the MCRA uncertainty limits with the exception of p70, where again the binning algorithm of DEEM may have caused additional deviations.

Table 3. Exposure US population to captan in apple (μg/kgBW/day). Percentiles MCRA and DEEM.

Percentage	MCRA Exposure	MCRA Lower Bound (p2.5)	MCRA Upper Bound (p97.5)	DEEM Exposure
60.00 %	0	0	0	0
70.00 %	0.0005497	3.277E-06	0.001238	0.003063
80.00 %	0.04209	0.02555	0.055	0.03905
90.00 %	0.4547	0.3518	0.5517	0.4458
95.00 %	1.512	1.2	1.919	1.521
97.50 %	3.782	2.801	4.564	3.672
99.00 %	8.727	6.424	10.38	8.451
99.50 %	14.42	10.26	16.99	14.03
99.75 %	22.11	15.49	26.63	21.69
99.90 %	33.4	24.7	41.83	34.32

4 Discussion

In this study the methods for probabilistic exposure assessment as described in the EFSA Guidance were implemented, and compared to alternative methods using a validation on simulated data. In pesticide monitoring the common pattern is to find a few positive values and a far greater number of non-detects. For statistical modelling this means that parametric distributions like the lognormal often have to rely on small numbers of observations. The problem is more severe in cumulative exposure assessment where also less important pesticides (with even fewer positive observations) are included in the CAG. The EFSA Guidance (EFSA 2012) has offered a proposal how a pessimistic cumulative assessment could be done. Our results show that the EFSA method may not work well. Due to all imputed LORs for NDs the cumulative distribution shows a strong spike and is not well modelled by a lognormal distribution. In our example the 99.9[th] percentile estimated by the 'pessimistic' approach was even lower than the one from the 'optimistic' approach. It may be concluded that further development of methodology is needed. This study presented some alternatives such as the censored lognormal model in combination with the use of additional data on agricultural use.

The results of the comparison between MCRA 8.0 and DEEM show in general a very good correspondence between the exposure percentiles calculated with the two programs. Extreme exposures in the upper tail of the distribution were almost identical. In the details there are some differences. Due to the binning system used for the Monte Carlo results in DEEM this program cannot give an accurate estimate of the left-hand side of the exposure distribution. MCRA on the other hand stores all Monte Carlo values and therefore is able to estimate the complete distribution. Another outcome of the comparison is that the percentages contribution per food are calculated differently. MCRA applies the sampling weights per individual of the consumption database when calculating the percentages contribution for each food, whereas DEEM calculates from the unweighted exposures.

Discussions are on-going about the proper methods for probabilistic risk assessment. On the one hand it is attractive to keep models as simple as possible, however this may lead to very unrealistic results, e.g. as was found in our case for the EFSA pessimistic model (see also Boon et al. in prep.). On the other hand, more realistic models may require data which are currently not always available. In this paper we showed that agricultural use data may add to the realism of the models, but such data should be made available from existing farmer registrations by national agencies, as for example given by Kennedy et al. (in prep.) for pesticide usage on crops in Great Britain.

It is likely that further harmonization of methodology for probabilistic risk assessment in Europe, America and the rest of the world will lead to further needs to be able to perform these assessments using available data from many different sources. Therefore the MCRA 8.0 model as developed in the ACROPOLIS project should not be seen as an endpoint, but as a step towards further integration.

Acknowledgements

References

- Boobis AR, Ossendorp BC, Banasiak U, Hamey PY, Sebestyen I, Moretto A (2008). Cumulative risk assessment of pesticide residues in food. *Toxicology Letters*, 180: 137-150.
- Boon, P.E., van der Voet, H., van Raaij, M.T.M. and van Klaveren, J.D. (2008). Cumulative risk assessment of the exposure to organophosphorus and carbamate insecticides in the Dutch diet. *Food and Chemical Toxicology*, 46: 3090-3098.
- Boon PE, van Donkersgoed G, Christodoulou D, Crépet A, D'Addezio L, Desvignes V, Ericsson BG, Galimberti F, Ioannou-Kakouri E, Hamborg Jensen B, Rehurkova I, Rety J, Ruprich J, Sand S, Stephenson C, Strömberg A, Turrini A, van der Voet H, Ziegler P, Hamey P, van Klaveren JD (in prep.). Cumulative dietary exposure to a selected group of pesticides of the triazole group in different European countries.
- Bosgra, S., van der Voet, H., Boon, P.E. and Slob, W. (2009). An integrated probabilistic framework for cumulative risk assessment of common mechanism chemicals in food: an example with organophosphorus pesticides. Regulatory Toxicology and Pharmacology. 54: 124–133.
- de Boer, W.J., van der Voet, H., Bokkers, B.G.H., Bakker, M.I. and Boon, P.E. (2009). A comparison of two models for estimating usual intake addressing zero consumptions and non-normality. *Food Additives and Contaminants. Part A – Chemistry Analysis Control Exposure & Risk Assessment,* 26: 1433-1449.
- Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, Tooze JA, Krebs-Smith SM. (2006). Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J Am Diet Assoc.* 106:1640–50.
- EC (2005). Regulation (EC) No 396/2005 of the European Parliament and of the Council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin and amending Council Directive 91/414/EEC. eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2005:070:0001:0016:en:PDF.
- EFSA (2008a). Opinion of the Scientific Panel on Plant Protection products and their Residues to evaluate the suitability of existing methodologies and, if appropriate, the identification of new approaches to assess cumulative and synergistic risks from pesticides to human health with a view to set MRLs for those pesticides in the frame of Regulation (EC) 396/2005. http://www.efsa.europa.eu/en/efsajournal/doc/705.pdf.

- EFSA (2008b). Scientific Opinion on Risk Assessment for a Selected Group of Pesticides from the Triazole Group to Test Possible Methodologies to Assess Cumulative Effects from Exposure through Food from these Pesticides on Human Health. http://www.efsa.europa.eu/en/efsajournal/doc/1167.pdf.
- EFSA (2010) European Food Safety Authority; Management of left-censored data in dietary exposure assessment of chemical substances. EFSA Journal 2010; 8(3):. [96 pp.]. doi:10.2903/j.efsa.2010.1557. Available online: www.efsa.europa.eu.
- EPA (2002). Guidance on Cumulative Risk Assessment of Pesticide Chemicals That Have a Common Mechanism of Toxicity. http://www.epa.gov/oppfead1/trac/science/cumulative_guidance.pdf.
- Gibney, M.J., van der Voet, H. (2003). Introduction to the Monte Carlo project and the approach to the validation of probabilistic models of dietary exposure to selected food chemicals. *Food Additives and Contaminants*, 20 (Suppl. 1): S1-S7.
- Goedhart PW, van der Voet H, Knüppel S, Dekkers ALM, Dodd, KW, Boeing H, van Klaveren JD (2012). A comparison by simulation of different methods to estimate the usual intake distribution for episodically consumed foods. Scientific Report submitted to EFSA. http://www.efsa.europa.eu/en/supporting/pub/299e.htm.
- Helsel, D.R. (2005). Nondetects and Data Analysis: Statistics for Censored Environmental Data. Wiley and Sons, New York.
- Kennedy M et al. (in prep.). A European model and case studies for aggregate exposure assessment of pesticides.
- Kennedy M et al. (in prep.). New approaches to uncertainty analysis for use in aggregate and cumulative risk assessment of pesticides,
- Meek, M.E., Boobis, A.R., Crofton, K.M., Heinemeyer, G., van Raaij, M., Vickers, C. (2011). Risk assessment of combined exposure to multiple chemicals: a WHO/ICPS framework. *Regulatory Toxicology and Pharmacology*, 60: S1-S14.
- Müller, A.K., Bosgra, S., Boon, P.E., van der Voet, H., Nielsen, E. and Ladefoged, O. (2009). Probabilistic cumulative risk assessment of anti-androgenic pesticides in food. *Food and Chemical Toxicology*, 47: 2951-2962.
- Petersen BJ (2010). Modeling dietary exposure with special sections on modeling aggregate and cumulative exposure. In: Hayes' Handbook of Pesticide Toxicology, Third Edition.
- Price PS, Chaisson CF (2005). A conceptual framework for modeling aggregate and cumulative exposures to chemicals. *Journal of Exposure Analysis and Environmental Epidemiology*, 15: 473-481.
- Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, Guenther PM, Carroll
- RJ, Kipnis V (2006). A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc*. 106:1575–87
- van der Voet, H., van der Heijden, G.W.A.M., Bos, P.M.J., Bosgra, S., Boon, P.E., Muri, S.D. and Brüschweiler, B.J. (2009). A model for probabilistic health impact assessment of exposure to food chemicals. *Food and Chemical Toxicology*, 47: 2926-2940.
- van Klaveren JD, Boon PE (2009). Probabilistic risk assessment of dietary exposure to single and multiple pesticide residues or contaminants: Summary of the work performed within the SAFE FOODS project. *Food and Chemical Toxicology*, 47: 2879-2882.
- van Klaveren JD, van Donkersgoed G, van der Voet H, Stephenson C, Boon PE (2009). Cumulative exposure assessment of triazole pesticides. Report 2009.008. RIKILT, Wageningen. Submitted to EFSA. http://www.efsa.europa.eu/en/supporting/doc/40e.pdf.