SZENT ISTVÁN UNIVERSITY

FACULTY OF FOOD SCIENCE

MODEL FOR FACILITATING THE PLANNING OF RISK-BASED MONITORING PROGRAMMES FOR PESTICIDE RESIDUES

Zsuzsanna Horváth

Gödöllő

2018

PhD School/Program

Name: PhD School of Food Sciences

Field: Food Sciences

Head of school: Gyula Vatai, PhD, DSc, Szent István University, Faculty of Food Sciences, Department of Food Engineering

Supervisor: Dr. habil. Ambrus Árpád, CSc, honorary professor, retired chief scientific advisor

Signature of acceptance of the Head of School and the Supervisor:

The applicant met all the requirements of the PhD regulation of Szent István University and the thesis is accepted for the defence process.

…...…… ……...

Head of PhD School Supervisor

1. INTRODUCTION

Pesticides are active substances which are designed to prevent, kill, weaken, attract or control any kind of pest. They can be used in the vegetation period of the plants, during storage, transportation or processing for foods and feeds as well. Use of pesticides contributes significantly to food security for the continuously augmenting population of the Earth. Their appropriate use can help the agriculture in providing enough good quality and quantity of foods and feeds. However, they are mostly toxic compounds, thus their non-compliant application can be hazardous for the human health and the environment too. Consequently, regulation of their use and monitoring of their appropriate application is needed. Crops intended for consumption should only contain residues of the pesticide in such quantity which does not cause harm to the human health. Monitoring of the commodities can be carried out by the competent authorities of the country and also by the producer. Nevertheless, it is practically impossible to check all pesticide residues in all products, as it would result a huge number of combinations.

Designing monitoring programmes for systematic control of the commodities is of paramount importance for the consumers, the market, the food producers and the authority as well.

2. OBJECTIVES

1. My main objective was to develop a model for assisting the planning of risk-based monitoring programmes for surveying the pesticide residues in raw agricultural commodities using all available relevant information.

2. In order to establish the theoretical basis and practical implementation of the model it was aimed to:

• Determine the parameters of the distribution of residues within- and between treated fields in individual crops and composite samples;

• Analyse the data from supervised field trials;

• Characterise the distribution of residue values in supervised field trial datasets;

• Investigate the impact of the factors affecting the uncertainty of the expectable maximum residue level (mrl);

• Consider the results of the historical monitoring data from preceding years and the ARfD (acute reference dose) values indicating the acute toxicity of pesticides;

• Investigate the relationship between the 97.5^{th} percentile (which is critical for the acute exposure assessment) and the median of the datasets.

3. In addition, my goal was to design a decision tree to facilitate the use of the model, and present practical examples for various situations that may occur in practice.

3. MATERIALS AND METHODS

3.1 Databases used for the analysis

3.1.1 Variability of residues in/on primary samples

Altogether 19600 residue values, derived from the following four different sources, were considered for the characterization of distribution of pesticide residues in primary samples:

1. Results of analysis of residues in medium sized fruits taken from 57 lots at wholesale distribution

points. A total of 90 to 110 fruits were taken from each lot and analysed with multi residue methods covering wide range of pesticides.

2. Specifically designed field studies including 53 sites and application of 25 different pesticide active ingredients representing the normal agricultural practice were conducted in Argentina, Austria, Brazil, Costa Rica, Croatia, Greece, Hungary, Italy, Malaysia, New Zealand, Philippines, Poland, Spain, and Thailand on small and large fruits and leafy vegetables.
3. Supervised field trials were also conducted by pesticide manufacturers coordinated by CropLife International on lettuce and grape as model crops. The results were evaluated and reported by the FAO/WHO Joint Meeting on Pesticide Residues.
4. In order to complement the available experimental data and to provide information on the distribution of residues in root vegetables and small-leafed vegetables, further field trials were conducted using carrot and parsley as model crops. The data base included 2637 residue data derived from 10 different pesticide active ingredients and 720 primary samples.

Only those residue – commodity combinations were taken into account where less than 10% of the results were below the limit of quantification (LOQ). In the latter cases the middle-bound approach was applied where the non-quantified values were considered as LOQ/2.

3.1.2 Field-to-field distribution of pesticide residues

A database comprised of residues from supervised field trials evaluated and reported by the Joint Meeting on Pesticide Residues (JMPR) between 1997 and 2011 gave the basis for the investigation of the field-to-field variation of pesticide residues. The database contained 25766 residues values of 1950 pesticide residue – commodity combinations.

For getting the residues of various magnitudes comparable, the individual residue values detected in supervised trials were normalized: the residue values were divided by the mean of their corresponding dataset. The normalized database (with $\mu = 1$) reflected the variability of the field-to-field distribution of residues in supervised field trials.

In order to identify the sources of variability of residues in individual data sets, a synthetic lognormal residue population consisting of 500000 data points with $\mu=1$ and SD=0.8 was generated with @Risk software (Version 5.0.0 Industrial Edition) for modelling the between fields distribution of pesticide residues. Descriptive statistical parameters of the normalized residues and the synthetic population are presented in Table 1.

With the exception of the minimum and maximum values, the descriptive statistical parameters of the normalized residue population and the synthetic database showed good agreement, which confirms the finding that the lognormal distribution describes well the distribution of residues.

Table 1. Descriptive statistical parameters of the normalized residues and the synthetic population

Parameters	Norm. residues μ=1; SD=0.794	Synthetic μ=1; SD=0.8
Min	0.00113	0.0311
Median	0.82260	0.7831
Average	1.00000	0.9988
Max	9.60084	18.8233
CV	0.79430	0.7915
P0.95	2.46530	2.4684
P0.975	3.00852	3.0773
P0.98	3.20000	3.2859
P0.99	3.97148	3.9693
n	25766	500000

The high residue values occurring above the 99th percentile in the databases practically did not have impact on our conclusions drawn from the results of the random sampling, as according to the current international practice the authorized maximum residue limit (MRL) has to cover the 95-99% of the residues in the composite samples while the acute exposure is calculated by taking into account the 97.5th percentile of the residues. Consequently, the random samples taken from the synthetic population realistically represent those from the supervised field trials.

3.2 Methods

3.2.1 Characterization of distribution of pesticide residues

The range of the residue values was characterised by their relative standard deviations also known as coefficient of variation (CV_R). The CV_R contains the sampling uncertainty

(CV_s) and the variability of the laboratory measurements (CV_L).

$$CV_R = \sqrt{CV_S^2 + CV_L^2} \quad (1)$$

Laboratories analyse the retained portions of homogenized samples (stored in freezer) according to the specification of ISO/IEC 17025 standard. These results provide information on the laboratory reproducibility (CV_L), including the heterogeneity of the homogenized sample.

The uncertainty of analyses of residues in crop units, CV_L, was generally much lower (15-20%) than the observed variability of residues in crop units and contributed only 1.5-2.5% to the CV_R values. Therefore, the uncertainty of analysed residues in crop units properly represents the within field variation of residues.

For the characterisation of the within and between field distributions of the pesticide residues the relative standard deviation (CV) of the datasets of composite samples and their weighed average (CV_{wt}) were taken into account.

The calculation of CV values of small residue datasets (≤10) was carried out with range statistics:

$$CV = \left(\frac{\sum \Delta}{n}\right) / d_2 \quad (2)$$

Where 'n' is the number of residue data points in one dataset, and d_2 is the corresponding factor of 1.128, 1.693, 2.059, 2.326, 2.534, 2.704, 2.847, 2.970 and 3.078 for n= 2, 3 ... and 10, respectively. Furthermore $\Delta = \frac{R_{max} - R_{min}}{\bar{R}}$, where R_{max}, R_{min} and $\bar{R}$ are the maximum, minimum and mean residue values in one dataset.

If the datasets consisted of larger than 10 values, the CV values were calculated from the standard deviation (SD) and mean value ($\bar{x}$) of residues:

$$SD = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}\,;\; CV = \frac{SD}{\bar{x}} \qquad (3)$$

The weighted average CV value for results of one commodity (CV_{wt}) was calculated as:

$$CV_{wt} = \frac{\sum(n \times CV)}{\sum n} \qquad (4)$$

While the weighted average for a commodity group (CV_{WT}) was also defined as:

$$CV_{WT} = \frac{\sum(k \times CV_{wt})}{\sum k} \qquad (5)$$

where k is the number of datasets in one commodity group.

Considering that the CV is underestimated in small residue datasets, the CV values calculated with the weighted average were modified with the Sokal correction as:

$$CV' = CV + \frac{1}{4 \times n} \qquad (6)$$

where CV' is the corrected CV and n is the number of data points which were used for calculation of the original CV value.

To compare the spread of residues in various datasets, the individual values were divided by the average of the dataset.

Residue distributions of primary samples were examined by using graphical representations in order to determine their typical characteristics. In addition, basic parametric distributions (lognormal, shifted lognormal, gamma, Weibull and normal) were fitted with @Risk

software on the combined normalised residue data in crop units. The goodness of the fit was characterised with the sum of squares of the relative differences of the input experimental data and the fitted distributions at the critical 5^{th} and 95^{th} percentiles. As food safety criterion, the compliance of 98% of commodities (βp=98%) with the MRLs at 95% probability (β_t) level was selected. The number of samples (n) to be tested was calculated with the following equation based on the binomial distribution:

$$\beta_t = 1 - \beta_p^n \text{ ; } n=\lg(1-\beta_t)/\lg(\beta_p) \qquad (7)$$

Based on the binomial theory, if we take 149 random samples and we found only one which contains residue above the MRL value, we can conclude with 95% of probability that the 98% of the products in the sampled lots comply with the regulation. If we cannot find any of the 149 samples which exceeds the MRL, the compliance level is even higher than 98%.

3.2.2 Modelling the uncertainty of mrl and HR estimation

For the determination of the variability of the highest residue (HR) and the expectable maximum residue level (mrl) random samples of 3, 5, 10, 25, 100, 120, 200 and 300 sizes were drawn 10000-10000 times with replacement from both the normalized residue database and from the synthetic generated population by a MS Excel VBA macro. The process was repeated three times for the sample sizes of 3 to 25. The data points in the samples represented the residue levels of a supervised field trial dataset. As next step, the average residue value, the CV of the residues, the ratio of the HR and median (M) value of each dataset and the corresponding mrls (using the OECD MRL calculator) were calculated.

3.2.3 Consideration of the Hungarian pesticide residue monitoring data

For assisting the prioritization of the pesticide residue – commodity combinations, a specific query format was elaborated for presenting the analytical results for each combination and other related parameters (e.g. LOQ, MRL) which were reported in the Hungarian laboratory system. Based on the results of the query, the software calculates weighting factors which can be considered in the model facilitating the planning of the risk-based monitoring programme.

4. RESULTS

Factors affecting the estimated acute exposure of consumers and the MRLs based on limited number of supervised trials were investigated in detail as they provide the basis for the first tier of the model assisting the planning of risk-based monitoring programme.

4.1 Characteristics of within field distribution of pesticide residues

Based on the 19600 residue values in the available database it was noted that the concentrations of residues measured in primary samples taken from a lot can be considered continuous in a hundredfold interval, where the high values appeared much less frequently but spread widely. The within field distribution of normalized residues in commodities of different sizes and shapes is very similar and can be described with a relative standard deviation of 80%, calculated from the weighted averages. Parametric distributions (lognormal, gamma, Weibull and normal) were fitted to the compiled dataset of the normalized residues of individual crops. The fitted Weibull, and gamma and

lognormal distributions overestimated the level of residues at the low concentration range, while they provided better approximation in higher ranges (Figure 1.). Goodness of fit was characterised with the sum of squares ($\Sigma\delta^2$) of relative differences of fitted and experimental distribution at the critical 95th percentile. It was concluded that the within field distribution of residues can be best described with a shifted lognormal distribution.

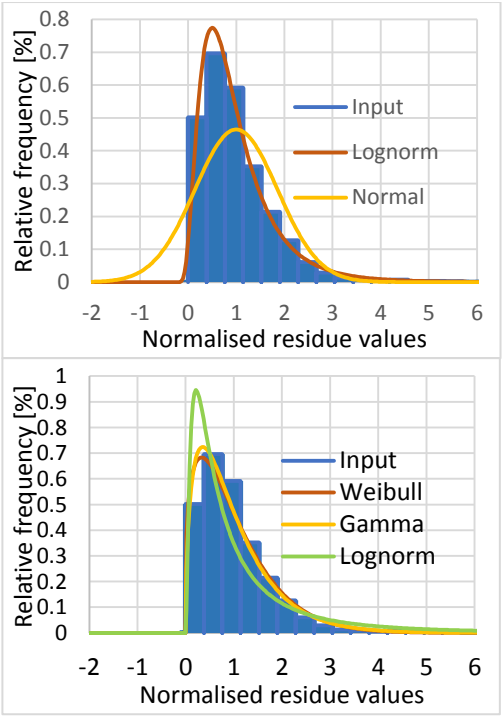

Figure 1. Examples for relative frequency distribution of the combined normalised residues in primary samples and fitted distributions (samples of leafy vegetables included lettuce, chicory, cabbage, parsley leaves)

CV_R ranges of different samples sizes taken from the database of primary samples of crops from one treated field and from the synthetic population were investigated. The following conclusions were drawn:

a) The range of minimum and maximum CV values, calculated from the results of repeated sampling, is very wide in case of small samples consisting of 5-10 residue values and underestimates the typical variability. Defining the distribution of residues based on such small samples is not recommended, as the result bears the large variability of the residues, thus their accuracy cannot be verified. By increasing the number of primary samples, the range of CV values is decreasing and gives better approximation to the true CV value of the parent population of the residues on the treated field. Average CV_R value of sample sizes over 300 differed less than 1,4% from the true CV of the synthetic parent population. Relative frequencies of CV values in different sample sizes are presented on Figure 2.

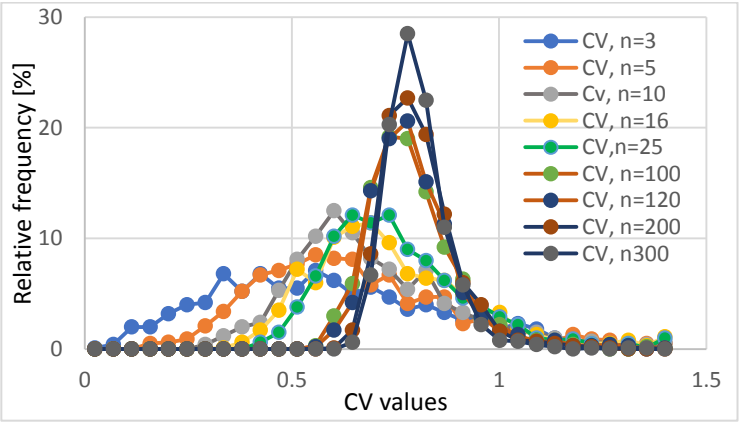

Figure 2. Relative frequency of CV values of datasets with n sample sizes drawn from the synthetic parent population

b) Maximum CV values (1.8, 2.1) of samples of 10-25 sizes drawn from the parent population cover the CV values of the most frequent sample sizes of supervised field trials.

c) CV values calculated from the residue values of random samples drawn 1000 times with sample sizes of 5, 10, 25, 100, 120 and 300 showed good linear relation (R^2=0,9998) with the expectable CV, which confirms that the general rule describing the relation of the composite and primary samples is valid even for strongly skewed populations:

$$V_n = \frac{V_1}{n} \; ; \; CV_n = \frac{CV_1}{\sqrt{n}} \qquad (8)$$

where V_1, CV_1 are the variance and the relative standard deviation of the data points of the population, and V_n, CV_n are the variance and relative standard deviation of the samples consisting of n data points.

d) Relative frequency distribution of CV values of samples with n sample size taken from a strongly skewed normalized primary residue population is slightly skewed towards the higher values, and the distribution is getting more symmetrical by increasing the sample sizes. In case of n≥25 the distribution is practically normal.

4.2 Characterization of distribution of pesticide residues between treated fields

Distribution of the 26766 residue values in the 1950 pesticide residue – commodity combinations taken from supervised residue trials reported in the JMPR evaluations were investigated. The number of samples in a dataset

varied between 5 and 121, while the most frequent sample sizes were 6 and 8.

The average residue level ranged from 0.0013 mg/kg to 712 mg/kg, whereas the ratio of the maximum and minimum residues (R_{max} / R_{min}) ranged from 21 (nuts) to 37961 (straw feed). The minimum and maximum CV values were between 2.1 (olives) and 21.5 (leafy vegetables).

Distribution of CV_R values of residues in supervised field datasets used for mrl estimation were examined. It was concluded that relative standard deviations of different crops and crop groups spread in a wide range, nevertheless their relative frequency distributions are very similar and overlapping, independently from their sample size.

Considering the weighted average of the CV_R values of 0.76 for the 1950 pesticide residue – commodity combinations and taking into account that the datasets most frequently contained 6 and 8 residue values, applying the Sokal correction resulted in 0.79 and 0.78 average CV_R values, respectively, which have a good agreement with the CV value of the parent population (0.794, see Table 1.). Consequently, the distribution of residues between treated fields can be well described with an average CV_R of 0.8.

The main conclusions drawn from the evaluation of the supervised field trial data were:

- no correlation was found between the CV of the residues in datasets and the pre-harvest intervals, or the number of trials;
- the effect of the non-quantified values (<LOQ) on the calculated CV value depends on the ratio of the average residue and the LOQ, and on the proportion of LOQ values in the dataset;

- the proportion of values below LOQ up to 50% had only small effect on the CV of residues;
- the expectable spread of the residues can be best described with the weighted average of their CV value;
- the difference between the weighted average CV of the commodity groups modified with the Sokal correction and the CV of the 25766 individual residues was only 1.4%, which indicates that the calculated CV values reflect well the real spread of the residues;
- CV ranges of different commodity groups were overlapping, and the variability of residues was similar regardless whether the residue data were combined from trials conducted in one or more countries, or in one or more commodity groups, and none of the factors could be identified as the major source of the variability of residues between the fields treated under practically the same conditions.

4.2.1 Factors affecting the uncertainty of estimation of the mrl and HR

The accuracy and uncertainty of estimation of the expectable maximum residue level (mrl) is affected (descending order of the magnitude of contribution) by the uncertainty of the measured residue values, the analytical method applied, the LOQ values and their proportion in the dataset, the range of the measured concentration of residues, especially the magnitude of HR value and the number of supervised residue trials.

4.2.1.1 Uncertainty of the measured residue values

The combined uncertainty of sampling and analysis (CV_R), ranging typically between 25-35% for composite samples

had the most pronounced effect on the residue values, which is especially relevant for the highest residue (HR) as it influences the mrl value, and the estimated acute exposure of the consumers.

4.2.1.2 Effect of non-quantified values (<LOQ) on the average and the CV value of the dataset

If the LOQ value is approximately the quarter of the average residue concentration of the dataset then it does not affect significantly the CV value of the dataset up to ~60% proportion of non-quantified values. While if the magnitude of the LOQ value is about the half of the average residue concentration then it already decreases the CV value of the dataset with nearly ~10% at 40% proportion. Considering that in our database the proportions of non-quantified values were less than 50% in the selected datasets and only 20.7% of the 25766 residues were <LOQ, the LOQ values in the dataset could only affect the average CV values to a small extent.

However, taking into account that the non-quantified values can affect the calculated mean, the median of the dataset was used as a reference point in our investigations.

4.2.1.3 Number of residue values in a dataset

The effect of the number of residue values in a dataset on the uncertainty of the estimated mrl and HR values was investigated with samples of different 'n' sample sizes (n=3, 5, 10 and 25) drawn by random sampling with replacement 10000-10000 times from the synthetic generated database. The sampling was repeated three times at each sample sizes.

CV values of the datasets ranged in a wide interval, but the range of the CV values were getting narrower by the increasing of the sample size as it was already concluded in

the case of the investigation of distribution of residues in primary samples (Figure 2.).

Samples were arranged in ascending order by their HR values. The 250^{th}, 5000^{th}, 9750^{th} datasets were selected and were transferred into the OECD MRL calculator where the mrl values were estimated. The selected datasets represent well the variability of residues and the expectable mrls in 95% of cases. The results indicate that even for sample sizes of 25, the HR value is underestimated in at least 50% of the cases. However, in over 50% of the cases, the calculated mrl values are more than three times higher than the 95^{th} percentile of the parent population. (The mrls should principally cover the 95^{th} and 99^{th} percentiles of the parent residue population.)

Investigating the relation of the 97.5th and 2.5th percentiles the ratios of the $HR_{0,975}/HR_{0,025}$ and $mrl_{0,975}/mrl_{0,025}$ were calculated for different sample sizes. The results are shown on Figure 3.

The average ratios of the maximum residues ($HR_{0.975}/HR_{0.025}$) and the $MRL_{0.975}/MRL_{0.025}$ calculated from the three sampling operations can be described with the following equations:

$$mrl_{0.975}/mrl_{0.025} = 11.197n^{-0.497};\ R^2 = 0.9931 \qquad (9)$$

$$HR_{0.975}/HR_{0.025} = 10.019n^{-0.292};\ R^2 = 0.9679 \qquad (10)$$

The $mrl_{0.975}/mrl_{0.025}$ and $HR_{0.975}/HR_{0.025}$ ratios describe well the uncertainty of values calculated from the supervised trials and confirm that for a reliable estimation sufficiently large number of supervised residue trials is needed.

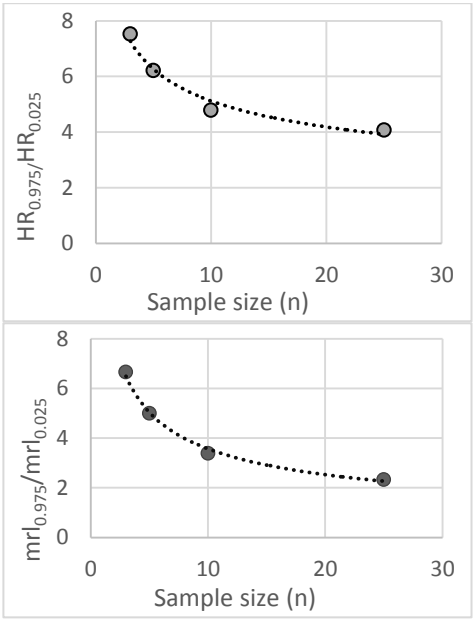

Figure 3. Ratios of HR and mrl values calculated at the 97.5th and the 2.5th percentiles at different sample sizes

Considering the influencing factors discussed above, it can be concluded that the results derived from one supervised field trial are only representing one possible residue level. If the residue trials were repeated with the same conditions (active substance, dose, DAL (days after last application)), the results obtained would be different from the ones from the first trial.

4.2.1.4 Ratio of HR and median residues ($F_{H/M}$)

The ratio of the HR and the median residues in the dataset can significantly affect the magnitude of the estimated mrl, thus the distribution of $F_{H/M}$ values was investigated in the 1950 datasets. The residue values were compared to the median (M, STMR) of the corresponding

dataset. The rounded cumulative % frequency distribution of the values in the datasets in the M≤R<3M; 3M≤R<4M; 4M≤R<5M; 5M≤R<6M; 6M≤R<7M, and R ≥ 7M median ranges were 54.5, 71.6, 78.6, 85.9, and 88.7 respectively, while in 11% of the cases the values were as high, or higher than sevenfold the median.

In case the HR value in a dataset is <3M then we can assume that the dataset in 45% of the cases does not correctly represent the true between fields differences of the residues, and based on these supervised field trials we underestimate the mrl and the exposure of the consumers.

4.2.1.5 The relation between the estimated mrl, the $F_{H/M}$ range and the sample size

Considering the importance of the most accurate estimation of the mrl value, further investigations were carried out for identifying the effects of the $F_{H/M}$ and the sample size on the estimation of the mrl and the exposure of consumers.

The analyses were made on the database introduced in section 3.1.2. For acquiring more information on the variability of mrl values, samples of different 'n' sample sizes (n=3, 5, 8, 10, 12, 16, 20 and 25) were drawn by random sampling with replacement 10000-10000 times with the MS Excel VBA macro. The datasets were transferred into the OECD MRL calculator with another macro, where the rounded mrl values were estimated and copied to the corresponding line of the database. As a reference point, the median value was used, and the parameters characterizing the spread of the residues were compared to the $F_{H/M}$ ratio. $F_{H/M}$ ratio was calculated for each sample, and the datasets were arranged in ascending order based on this parameter. CV value was calculated for the mrl and HR values and

further characteristic parameters for the distribution of mrl values were defined for the following $F_{H/M}$ ranges:

$F_{H/M}$ <3; 3≤ $F_{H/M}$ <4; 4≤ $F_{H/M}$ <5; 5≤ $F_{H/M}$ <6; 6≤ $F_{H/M}$ <7; $F_{H/M}$≥7.

Regardless the sample size, the majority of the generated datasets were in the $F_{H/M}$ <3 category. The range of the calculated mrls depends on the sample size (n) of the dataset, and on magnitude of the $F_{H/M}$ range. As it is presented on Figure 4., the bigger is the sample size the smaller is the difference between the minimum and maximum estimated mrls in the $F_{H/M}$ ranges which indicates more precisely estimated mrl and exposure.

In the range of $F_{H/M}$ <3 - which results in the most uncertain estimation – at samples sizes of 3 and 5 the $\delta_{MRL}=mrl_{max}/mrl_{min}$ ratio indicating the uncertainty of mrl estimation were 67 and 25, respectively, which was considered unacceptably high.

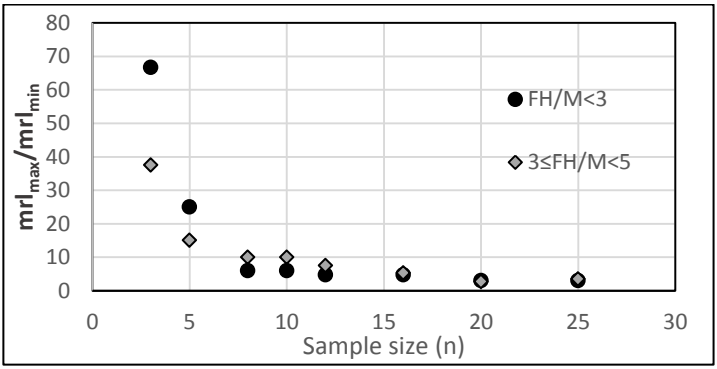

Figure 4. The relation of the calculated mrl_{max}/mrl_{min} ratio and the sample size (n) of the datasets in the $F_{H/M}$<3 and 3≤$F_{H/M}$<5 ranges.

For the accurate estimation of the MRL and the risk assessment of consumers, minimum 8, ideally 16-25

appropriately implemented supervised residue trials are needed resulting δ_{MRL} values of 6, 4 and 3, while the ratio of underestimated mrls at a sample size of ≥16 is only around 1%. (Figure 5.)

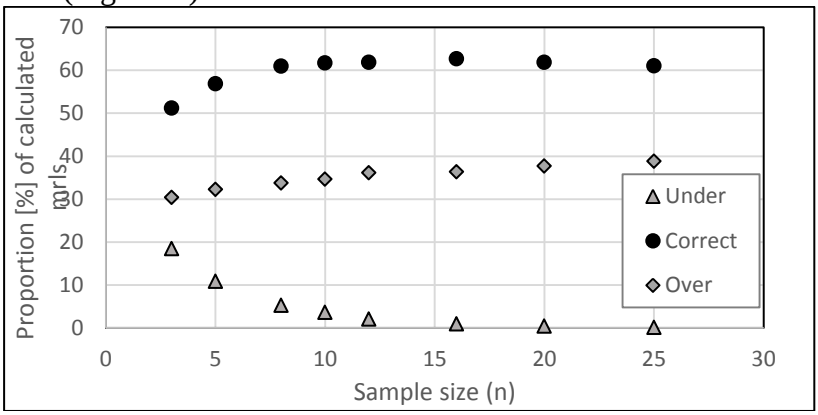

Figure 5. Percentage of under, correctly and over-estimated mrls as a function of sample size in $F_{H/M}$ range <7.

4.2.2 Relation of the HR and median

Results presented in section 4.2.1.5 indicated that if the $F_{H/M}$ ratio is ≤ 3 or ≤ 4 then there is a high probability that the expectable maximum residue level is underestimated. Thus, the relation of the input parameters of the exposure assessment of consumers (97.5^{th} percentile of the residues and the median) and the sample size was further investigated. Sample sizes of n=4 to 32 were drawn with replacement 10000 times from the database of the 25766 normalized residue values of supervised field trials. The distribution of the normalized residues, the HRs, and medians of the datasets for sample size 8 (the most frequently occurring sample size in the JMPR evaluations) are presented on Figure 6.

To estimate the HR value expectable in 95% of the cases from the limited number of available residue values,

the 5th percentiles of the medians of the samples were compared to the known 97.5th percentile of the parent population.

$$f_{M,n} = \frac{3{,}009}{P0{,}05_{M,n}} \qquad (11)$$

The $f_{M,n}$ was calculated with equation 11. for all sample populations of different sizes. The ratio gives, with 95% probability, the 97.5th percentile of residues relative to the median of the corresponding dataset derived from the application of pesticide according to the experimental conditions.

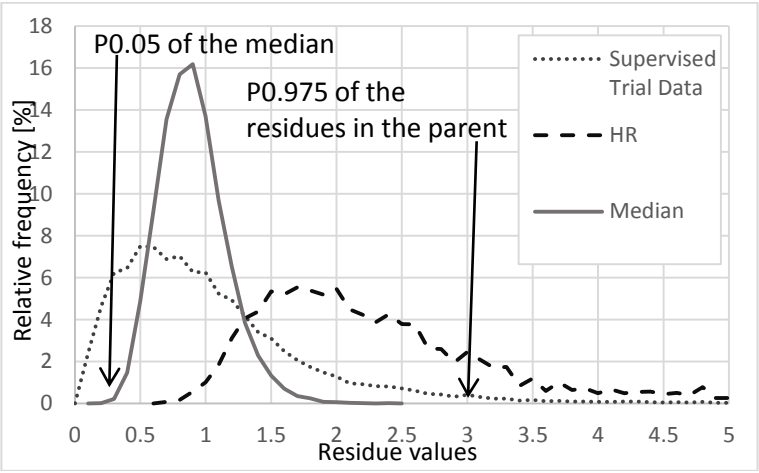

Figure 6. Distribution of normalized residues, and the median and HR values of residues in replicate samples of size 8 drawn from the parent population (25766) of normalised supervised trial data.

The relationship of $f_{M,n}$ and the size of random samples drawn from the normalized parent population is described with equation 12.

$$f_{M,n} = 10.233 \times n^{-0.228} \quad (R^2=0.9909) \qquad (12)$$

The expectable highest residue $HR_{P0.975}$ in the 95% of the cases is calculated as:

$$HR_{P0.975} = f_{M,n} \times STMR \qquad (13)$$

Factors of $f_{M,n,0,975}$ calculated with equation 13. at sample sizes 3 to 16 are presented in table 2. The factors are approximately reflecting the distribution of residues which shows that (at sample sizes of at least 5) the 89% of the residues was in the <7M range.

The spread of the residue values at sample sizes of 4 or less is larger, and can be described with a factor of 7.5 or 8.

Table 2. Factors for calculation of expectable highest HR values with 95% probability

n	$f_{M,n,0.975}$	n	$f_{M,n,0.975}$
3	8.0	10	6.1
4	7.5	11	5.9
5	7.1	12	5.8
6	6.8	13	5.7
7	6.6	14	5.6
8	6.4	15	5.5
9	6.2	16	5.4

4.3 Model for facilitating the planning of risk-based monitoring programmes

Based on the results of the supervised field trials and the modelling of residue distributions, a two-tiered model was elaborated for the prioritization of the different pesticide residue – commodity combinations, which provides gradually more accurate results depending on the information available.

1st Tier

The first tier uses the data from supervised trials conducted before the registration of a pesticide. Two different factors are calculated at this stage.

- F_{MRL} reflects the uncertainty of estimation of the maximum residue level based on supervised trial data.
- F_{ast} indicates the risk from the potential acute exposure in relation to the ARfD. (if defined)

The weighting factor, used for ranking the commodity–pesticide residue combination, F_{TI}, is the higher of the F_{MRL} or F_{ast}.

(a) Calculation of F_{MRL}

$$F_{MRL} = f_{ST} + f_{n\beta p} \quad (14)$$

The f_{ST} is calculated from the ratio of the MRL and STMR based on the cumulative frequency of residues in the median ranges.

(b) $$f_{ST} = 100 - \Sigma P\% \quad (15)$$

The corresponding factors are shown in table 3.

Table 3. The f_{ST} values in the median ranges

MRL/STMR ratio	ΣP%	f_{ST}
>7 M	100	0
6M≤ MRL<7M	89	11
5M≤ MRL <6M	86	14
4M≤ MRL <5M	79	21
3M≤ MRL <4M	72	28
<3 M	55	45

The $f_{n\beta p}$ factor considers the uncertainty of estimation of MRLs as a result of limited number of supervised trials:

(c) $$f_{n\beta p} = 0.5(100 - \beta_t\%) \quad (16)$$

where β_t is calculated with equation 7. from the sample size of supervised trials (n) assuming that 95% of the residues (βp=0.95) are complying with the MRL according to the underlying principles of OECD MRL calculator. An adjusting factor of 0.5 is applied to give proportional weight for the factors of f_{ST} and $f_{n\beta p}$.

(d) Calculation of F_{ast} factor for the consideration of acute risk to the consumers.
It is calculated only if ARfD has been established.

$$F_{ast} = \frac{ESTI_e}{ARfD}\% \qquad (17)$$

The $ESTI_e$ is calculated with the $HR_{P0.975}$ value obtained with equation 13. The acute exposure calculated with the likely highest residue is indicated with $ESTI_e$ to be distinguished from the one, calculated with the usual method, using the HR of the supervised field trials (ESTI). The WHO's 'Template for the evaluation of the acute exposure' or the EFSA PRIMo3 model can be applied for the calculation of $ESTI_e$, after inserting in the template the $HR_{P0.975}$ and the corresponding national food consumption data of the given commodity, if available. In other cases, the intake figures included in the template are used.
The template gives the % exceedance of the maximum ARfD which is equal to F_{ast}. factor.

2nd Tier

Based on the monitoring data of the preceding years, the specific query format (see section 3.2.3) calculates a weighing factor (F_{M0}) which indicates the priority of the given pesticide residue-commodity combination. It is calculated by taking into account the combined effect of the number of residue data (N) derived from monitoring

programmes and the frequency (f_p) of occurrence of residues.

$$F_{M0} = (f_m + f_p) \quad (18)$$

Following the recommendation of the BASLINE project we aim to test the compliance of ≥ 98% of the marketed commodities (β_p=0.98) with the established MRL and determine the probability of occurrence of residues (β_t) exceeding the MRL with equation 7.

The f_m weighting factor is calculated as:

$$f_m=100*(1-\beta_t) =100*\beta_p{}^N \quad (19)$$

Frequency of occurrence of detectable residues is taken into account by f_p:

$$f_p = 100\frac{\sum_{i=1}^{n} R_i \times MRL^{-1}}{N} \quad (20)$$

Where R_i–s are the measured residues in the samples.

The potential for acute intake problem based on the ratio of $ESTI_M$ and ARfD is considered by F_{aM}:

$$F_{aM} = \frac{ESTI_M}{ARfD} \quad (21)$$

The $ESTI_M$ is calculated from the largest residue detected in samples derived from monitoring programmes taking into account the definition of residues for dietary intake calculations.

The weighting factor, used for ranking the commodity–pesticide residue combination, F_{T2}, is the higher of the F_{M0} or F_{aM}.

4.3.2 Application of the risk-based tiered model

A decision tree, shown in figure 7., was elaborated for assisting planning monitoring programmes and application of the model in different situations.

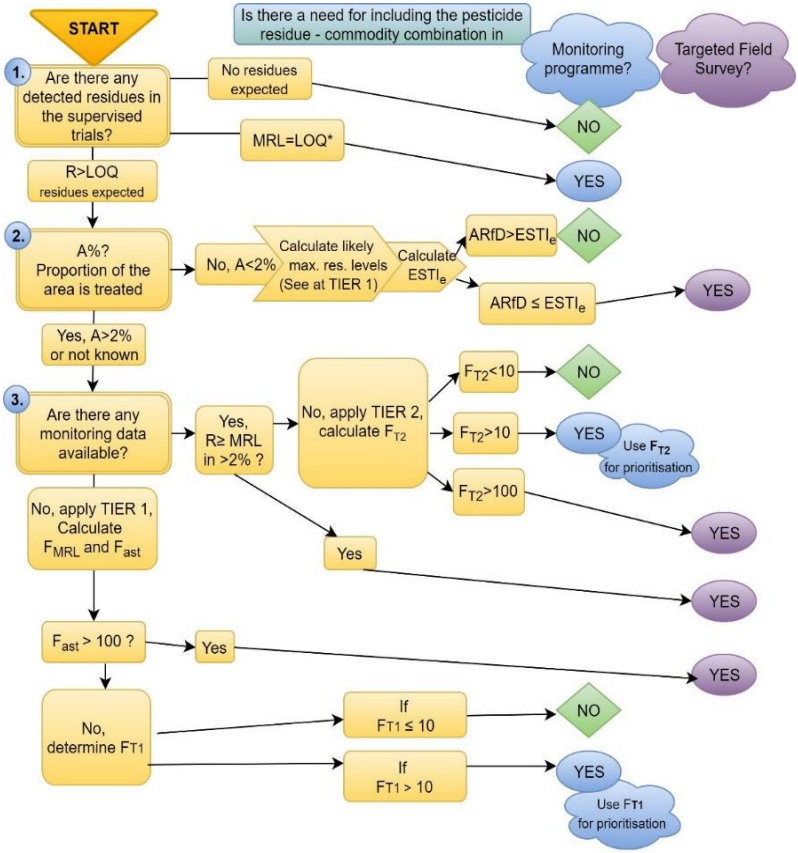

Figure 7. Decision tree for the application of tiered ranking model.

ESTIe: estimated short term intake with 95% probability level, A%: percentage of area expected to be treated with a given pesticide from the total cultivated area of the commodity; F_{MRL}, F_{M0}: weighting factors; F_{aM}, F_{ast}: factors indicating short term intake concern. F_{T1}: weighting factor to be used at the first tier of the model, the higher from F_{MRL} and $F_{ast;}$ F_{T2}: weighting factor to be used at the second tier of the model, the higher from F_{aM} and F_{M0}

Step 1: Evaluation of the information available for the registration of the pesticide.

Review the results of supervised residue trials:

1. Would detectable residue remain in the harvested crop?
   - 1.1 No detectable residue is expected in or on treated crop → no need to test the pesticide residue – commodity combination in the monitoring programme.
   - 1.2 MRL=LOQ*, this situation may occur when:
     - no detectable residue can be expected with analytical methods available at the time of the registration of the pesticide, used according to the label instruction;
     - the registration authority does not want the compound to be used, therefore the MRL is set at LOQ. (In these cases, an MRL=0.01 mg/kg is usually set in the EU) Residues may occur in crops if the product is illegally used or in commodities imported from countries where the product is authorized. Therefore, the commodities concerned may be included in the monitoring programme.
   - 1.3 R>LOQ, residues are expected: The use of compound results in detectable residues in treated commodities → go to step 2.

Step 2: The commodities are evaluated which can be treated based on the authorization. The potentially treated area expressed as percentage (A%) of total growing area of the given crop is taken into account:

- 2.1 A%? Is the percentage of the area treated with a given pesticide from the total cultivated area of the commodity known? How big is this area?

- If treated area < 2%, calculate the $HR_{P0.975}$ based on principles described in the 1st Tier and estimate the short-term intake ($ESTI_e$) with equation 13.
  - In this case, if the pesticide has no ARfD, or $ESTI_e$ < ARfD we can consider that there is no acute risk, thus there is no need to include the pesticide in the monitoring programme (or only in very exceptional cases).
- If $ESTI_e \geq$ ARfD → initiate targeted field surveys, because finding treated crops would have very low probability in case of random sampling of all marketed commodities, but the acute risk is considerable;
- If the treated area (A%) is larger than 2% or not known →go to step 3.

Step 3: historical monitoring data are considered.

Are results of previous monitoring programmes available?

3.1 Yes, we use the 2nd Tier of the model and consider the available historical monitoring data.

- Is the proportion of R ≥ MRL larger than 2%?
  - If yes, targeted field survey is recommended;
  - If no, apply 2nd Tier of the model and calculate combined weighing factor (F_{M0}) and factor reflecting the acute risk (F_{aM}), if there is ARfD for the pesticide. Choose the larger one as F_{T2}.
    - $F_{T2} \leq 10$ → the pesticide residue – commodity combination does not need to be tested;
    - $F_{T2} > 10$ → include the pesticide-commodity combination in random monitoring programme with the weight given by the factor
    - $F_{T2} \geq 100$ → initiate targeted field survey.

3.2 No monitoring data are available.

- apply 1st Tier of the model and evaluate supervised trial residue data. Calculate F_{MRL} and F_{ast} (if applicable). Choose the higher as F_{T1}.
  - If Fast $\geq$ 100 $\rightarrow$ initiate targeted field surveys, otherwise:
    - F_{T1} < 10? $\rightarrow$ the pesticide residue – commodity combination does not need to be tested;
    - $F_{T1} \geq$ 10? $\rightarrow$ rank the commodity-pesticide combination for inclusion in random monitoring programme based on F_{T1}.

Seven examples, including points to note, are given in the dissertation for the application of the model in different possibly occurring situations. However, they are not presented here due to space limitations.

4.4 Determination of number of samples to be included in future monitoring programmes

Optimally, for each commodity at least 149 random samples should be analysed over one or more growing seasons. This would enable finding defective lots containing residues above the MRL over 2% of cases with 95% probability. Regarding that the thousands of products marketed cannot be monitored with such frequency, if the risk is lower, the non-compliant lots could be identified with lower probability by analysing fewer samples.

Ideally the number of samples to be included in the monitoring programme can be determined based on the calculated weighing factor (Table 4.) Where the sufficient testing capacity or financial resources are not available, the most critical commodity pesticide combinations should be

given the priority and allocating the number of samples to other commodities proportionally to their weight.

Table 4. Recommended number of samples depending on the calculated weighting factor (F).

F	N	βt%[1]
≥100	149	95
≥75	114	90
≥50	94	85
≥40	60	70
≥30	46	60
≥20	30	45
≥15	15	25
≥10	10	18
<10	0	0

[1]: probability of detection of the non-compliant lot assuming 98% compliance

For clarifying specific problematic situations targeted field surveys are recommended to be initiated, where sampling of treated commodities should also be based on random selection of sites. Based on the experiences gained in the investigation of sampling uncertainty, it is recommended to take at least ≥2 individual replicate samples from ≥ 8 treated fields reflecting the recommended application conditions. Sampling of more than 20 areas practically does not affect the reliability of the results.

4.5 New scientific results

19600 residue values were evaluated, derived from 90-320 individual units of 24 different commodities taken from 113 independent fields, treated with 47 pesticides according to the normal agricultural practice in 27 countries of five

continents. The results were completed with modelling of synthetic lognormal database consisting of 500000 data points, generated based on the parameters of the experimental residue population.

1. *Based on the results of the residue analyses and the modelling experiments, the within field distribution of residues from primary samples were characterized. It was also presented that no reliable conclusion can be drawn regarding the range of the expectable residues on the basis of the analyses of 5-25 samples, as minimally 100-100 random samples from 8-20 fields should be taken for this investigation.*

The between fields variability of residue concentrations was investigated on the basis of the supervised trial data used by the FAO/WHO Joint Meeting on Pesticide Residues for the estimation of mrls, which consisted of 25766 residue values (with typically 10-25 primary samples in one composite sample) of 1950 pesticide residue – commodity combinations. The number of samples in a dataset varied between 5 and 121, while most frequently 6 and 8 supervised trials made up one dataset. For getting the spread of residues in datasets with largely varying average residues comparable, the individual residue values were normalised by dividing them with their average to obtain datasets with a mean of 1. The compiled dataset consisted of 25766 results with an average value of 1 and a CV_{ah} value of 0.794. From this database synthetic supervised trial detests consisting of different number of residue values were generated with random sampling with replacement in order to investigate the factors affecting the field-to-field variation of residues. In the datasets, the spread of the residues was characterized with the ratio of the highest residue and the median ($F_{H/M}$)

as the median value is not affected by the LOQ values if they are below 50%.

2. *On the basis of the supervised field trial database and the modelling experiments, the field-to-field variability of the average residue concentrations, the range of the residues in supervised field trial datasets relative to their median, the factors affecting the uncertainty of the mrl estimation, the minimum and optimal numbers of supervised trials to be used for mrl estimation were determined. In order to estimate the mrl and the consumer exposure with an acceptable accuracy, minimum 8 optimally 16-25 supervised residue trials are necessary regardless of the proportion of the treated field to the whole cultivated area, This finding is in contrast to the current Codex guidelines requiring only 4-5 supervised field trial residues as a minimum.*

The relation of the input parameters of the exposure assessment of consumers (97.5^{th} percentile of the residues and the median) and the sample size was investigated to facilitate the possibly most accurate estimation of the exposure of consumers. For this reason, the 5^{th} percentiles of the medians of the datasets consisting of 4 to 32 samples were compared to the known 97.5^{th} percentile of the parent population (3.009). By this way the HR was estimated which would cover the true 97.5^{th} percentile of resides with 95% probability in case of various number of trials. The ratios of the true 97.5^{th} percentile of residues and the 5^{th} percentile of the median of the datasets were calculated for each dataset consisting of 4-32 samples: $f_{M,n} = \frac{3.009}{P0.05_{M,n}}$

3. *The ratio of the known 97,5th percentile of the parent population and the median of the 5th percentile of the*

datasets ($f_{M,n}$) depends on the sample size (n), which can be described with the following equation:

$$f_{M,n} = 10.233 \times n^{-0.228} \quad (R^2=0.9909)$$

The highest HR foreseen in 95% of the cases can be calculated from the supervised trial median residue (STMR) as:

$$HR_{P0.975} = f_{M,n} \times STMR$$

The STMR values defined for exposure assessment should be inserted in the equation. If it is not available, the results of the monitoring can be used with a corresponding adjusting factor.

The coherence of MRL based on limited number of supervised field trials and the practical use of the pesticide as well as the exposure of consumers derived from the practical use of the pesticides should be regularly monitored. As the several thousands of pesticide residue – commodity combinations cannot be equally monitored in practice (due to the laboratory capacity and financial constraints), they need to be prioritized. For the risk-based ranking of the commodities, no quantitative method is currently available.

4. *A two-stage model for planning monitoring programmes was elaborated, which ranks the pesticide-commodity combinations based on weighting factors considering:*
    - *all information available at the time of the registration of the pesticide at the first stage,*
    - *the historical monitoring data from the preceding years at the second stage.*

Application of the model was demonstrated with practical examples of different possibly occurring situations, and a method for calculation of sample size was recommended for testing the compliance of marketed commodities with a 95% of probability of detection of non-compliant lot in random monitoring or in targeted field surveys.

5. CONCLUSIONS AND RECOMMENDATIONS

The database consisting of 19600 results of pesticide residue analyses was compiled from samples coming from fields cultivated at different technological levels, considered sufficiently large for the evaluation of the nature of the distribution of residues in individual crop units (primary samples) within a treated field. The effects of different factors influencing the variability of residue values (such as the characteristics of plants, the physical-chemical properties of plant protection products, the application method, the weather) are cannot be distinguished based on 100-120 samples from one site. Consequently, recommendations on the distribution of residues concerning small subgroups or individual commodities cannot be made, as sufficiently large residue databases could not be established due to high costs.

Specifically, for the determination of the characteristics of distribution of one selected commodity – pesticide residue combination minimally 100-100 random samples from 8-20 fields should be taken. No reliable conclusion can be drawn regarding the range of the expectable residues based on the analyses of 5-25 samples taken from 1-2 fields, as the variability of residues cannot be properly taken into account on such basis.

As such an ideal database is not available, based on the 19600 residue values in primary samples an average within field distribution of residues is recommended to be described with a relative standard deviation of 80 %.
Due to the wide range of residue values on the treated areas, the inevitable sampling uncertainty substantially influence the average residue content of composite samples which has to be considered while the sources of uncertainty of the mrl estimation are investigated.

The supervised trial data used by the FAO/WHO JMPR for estimation of mrls were extracted from the annually published reports consisted of 25766 residue values of 1950 pesticide residue – commodity combinations and reflected the agricultural practices of five continents, thus they provided a good basis for the evaluation of the nature of the field-to-field distribution of residues and the estimation expectable maximum level.

It has to be noted that the supervised field trials are conducted under strictly controlled conditions, with optimal application technology, in small production area. Thus, under practical conditions the variability of the residues can be larger than the one determined from supervised field trials.

To investigate the field-to-field variation of pesticide residues, the relative standard deviation (CV) was calculated which allowed the comparison of the variability of residues in datasets with different residue concentrations.

The variability of the supervised residue database (~ 79%) was calculated based on the weighted average of the CV values of the datasets. It was concluded that the CV values had wide ranges and the variability of different commodity – pesticides residue combinations were covering

each other. None of the factors could be identified as the major source of the variability of residues between the fields treated under practically the same conditions.

Consequently, the compiled database, containing the normalized residues of the datasets could be used for the analysis of relations, and can be used for drawing general conclusions.

The value of the expectable maximum residue level (mrl) is affected by the uncertainty of the measured residue values, the range of the measured concentration of residues, especially the magnitude of median and HR value and the number of supervised residue trials. Taking into account that also the LOQ values and their proportion can affect the calculated mean, the median value was used as a reference point in the investigations.

The residue values of datasets were arranged into the median ranges (M, STMR) of the corresponding dataset. While only 54% of the experimental data were <3M, the 88,7% of the cases were within the <7M range which indicates that during the practical application of the pesticide, a significant proportion of the residues are above the 3M range. Consequently, the current principles of the international practice and the Codex guidelines, which only requires 4-5 residue values for the determination of the mrl values, do not give a strong basis for the accurate estimation of the maximum residue limit (MRL). The underestimated mrl (and MRL) gives a false picture about the true exposure of consumers and brings the risk that the food producer will be fined, or the commodities have to be repealed due to their high residue content in spite of the fact that the use recommendations were followed during the production of the commodity. On the other hand, the overestimated MRL

allows higher dosage of application than it is needed, which leads to unnecessarily higher exposure of the consumers.

For the protection of the consumers and the safe marketing of the plant commodities, it is recommended to conduct minimum 8, ideally 16-25 appropriately designed supervised trials, regardless the area of cultivation of the food commodity. When the number of supervised residue trials is defined, it should be considered that the marketed products has to comply with the MRL if we want to avoid the loose of trust between the farmers and traders, severe economic loss and obstacles in the export-import of the product. The national competent authorities can initiate further supervised trials in addition to the ones carried out by the pesticide producing companies. Considering that the planning and method development costs of the supervised trials are not increased by increasing the number of trials, the costs of the additional trials are not significant comparing to the value of the marketed products. Furthermore, the costs of the additional trials can be shared between the producer, the exporters and the wholesalers.

The elaborated risk-based monitoring model – assuming the compliance of the pesticide applications with the GAP – takes into account the probability of non-compliant lots or short-term intake exceeding the ARfD values by using the data from supervised trials conducted before the registration of a pesticide or residue data derived from monitoring programmes of the preceding years for the prioritization of different pesticide residue – commodity combinations.

For the further refinements of the priorities defined by the model, country- / cultivation area-specific factors can be and should be taken into account, such as the importance of

the given products in the economy/export revenue of the country, the sources of purchases of the plant protection products and the most frequently marketed and used pesticides in the country.

The model cannot be used automatically, thus the planning of the monitoring programmes is not the task of one person but it should be preferably carried out by a team of competent experts who have the necessary knowledge and information.

At the 1st Tier of the model special attention should be paid during the determination of the priorities for the followings:

- *when the number of supervised trials used for establishing the MRL is less than 8 and or the highest residue value is within the range of 3M-4M, there is a high probability that the MRL and the exposure of consumers are underestimated;*
- *pesticides or their metabolites which can only be analysed by single residue methods, as most of the laboratories determine them only in specified cases, as additional time and cost are involved;*
- *commodities covered by group-MRL (e.g. leafy vegetables) where separate supervised field trials for each crop were not carried out.*

The commodities included in the monitoring programmes should be analysed with multi-residue (MRM) methods enabling the simultaneous determination of as many residues as possible, thus additional information can be collected about the concentration and frequency of occurrence of pesticide residues which can be used in the planning of future monitoring programmes.

Considering that determination of the compliance of a commodity with 100% probability is an unreachable goal in practice, as food safety and production safety criteria, a compliance level of 98% was recommended by the BASELINE consortium, which gives a good starting point for the planning of monitoring programmes. In those cases, where there is a possibility for the acute exposure of consumers exceeds the ARfD or noncompliant samples are occurring over 2% of the marketed lots, the above criteria should be verified with 95% of probability by taking and analysing at least 149 random samples. In other cases, based on the available sampling and laboratory capacity and the financial conditions, the number of samples can be decreased which certainly results in the decrease of probability of detection of residues at or above the MRL.

Reason of the violation of the MRL as a consequence of the application of the authorized plant protection technology can be more effectively identified through targeted monitoring implemented by taking duplicate samples from randomly selected 8-20 fields with known pesticide treatment history.

6. PUBLICATIONS

Peer reviewed publications in the subject of the dissertation

Zsuzsanna Horváth, Árpád Ambrus, László Mészáros & Simone Braun (2013) Characterization of distribution of pesticide residues in crop units, Journal of Environmental Science and Health, Part B: Pesticides, Food Contaminants, and Agricultural Wastes, 48:8, 615-625 / IF: 1.362

Zsuzsa Farkas, Zsuzsanna Horváth, Kata Kerekes, Árpád Ambrus, András Hámos & Mária Szeitzné Szabó (2014) Estimation of sampling uncertainty for pesticide residues in root vegetable crops, Journal of Environmental Science and Health, Part B: Pesticides, Food Contaminants, and Agricultural Wastes, 49:1, 1-14. / IF: 1.362

Zsuzsanna Horváth, Judit Sali, Andrea Zentai, Enikő Dorogházi, Zsuzsa Farkas, Kata Kerekes & Árpád Ambrus (2014) Limitations in the determination of maximum residue limits and highest residues of pesticides: Part I, Journal of Environmental Science and Health, Part B: Pesticides, Food Contaminants, and Agricultural Wastes, 49:3, 143-152 / IF: 1.362

Árpád Ambrus, Zsuzsanna Horváth, Zsuzsa Farkas, István J. Szabó, Enikő Dorogházi & Mária Szeitzné-Szabó (2014) Nature of the field-to-field distribution of pesticide residues; Journal of Environmental Science and Health, Part B: Pesticides, Food Contaminants, and Agricultural Wastes, 49:4, 229-244 / IF: 1.362

Zsuzsa Farkas, Zsuzsanna Horváth, István J. Szabó, and Árpád Ambrus (2015) Estimation of sampling uncertainty of pesticide residues based on supervised residue trial data; Journal of Agricultural and Food Chemistry., 63 (18), pp 4409–4417 / IF: 3.154

Árpád Ambrus, Zsuzsanna Horváth, Júlia Szenczi-Cseh & István J. Szabó (2018). Factors affecting the quantitative uncertainty of the estimated short-term intake. Part I — Calculation methods. Journal of Environmental Science and Health - Part B Pesticides, Food Contaminants, and Agricultural Wastes, 53(6), 394-403. / IF: 1.362

Árpád Ambrus, Zsuzsanna Horváth, Júlia Szenczi-Cseh (2018). Factors affecting the quantitative uncertainty of the estimated short-term intake. Part II — Practical examples. Journal of Environmental Science and Health - Part B Pesticides, Food Contaminants, and Agricultural Wastes, 53(6), 404-410. / IF: 1.362

Book chapter

Zsuzsanna Horváth és Árpád Ambrus (2017) Principles of Control of Small-Scale Production of Fruits and Vegetables and Planning Risk-based Monitoring Programmes, in Ambrus Á. Hamilton D. (szerk.) Food Safety Assessment of Pesticide Residues, World Scientific Publishing Europe Ltd.: London 467-506.

Not peer reviewed publications

Árpád Ambrus, Zsuzsa Farkas, Zsuzsanna Horváth, Gabriella Kötelesné Suszter: Principles and practices of control of pesticide residues in food, Journal of Food Investigation LX, 2, 8-32 2014

Conference abstracts, summaries

Zsuzsanna Horváth, István Ficzere, Árpád Ambrus: Növényvédőszer-maradékok eloszlásának vizsgálata egyedi terményekben, „Fiatal kutatók az egészséges élelmiszerekért" konferencia, Debrecen, 19th February 2013

Zsuzsanna Horváth, Árpád Ambrus: Principles for planning monitoring pesticide residues in agricultural commodities, 4th MoniQA International Conference, Budapest, 26th February – 1st March 2013

Zsuzsanna Horváth, Árpád Ambrus: Establishing performance criteria for testing compliance of chemical contaminants with legal limits (BASELINE Final Conference, Bologna, Italy, 12th November 2013)