Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues

Árpád Ambrus* and Júlia Szenczi-Cseh

1Retired Scientific Adviser of National Food Chain Safety Office, Hungary
2National Food Chain Safety Office, Directorate for Food Safety Risk Assessment, Budapest, Hungary

*Corresponding Author: Árpád Ambrus, Retired Scientific Adviser of National Food Chain Safety Office, Hungary.

Received: February 23, 2017; Published: March 20, 2017

Abstract

Pesticides are generally toxic chemicals. Before their use is authorized, the exposure of consumers to their residues remaining in the food is usually assessed with methods providing point estimate. The calculation is simple if all information is available, but it is rarely the case. The risk managers should be aware of the uncertainties associated with the calculated values to make the right decision. The nature and magnitude of uncertainties of numerous factors affecting the combined uncertainty of the reported exposure are reviewed in this paper. Further on, guidance is given on the utilisation of available information for making expert judgement to obtain the best estimate for replacing the missing information. The uncertainties of parameters influencing the calculated dietary exposure vary at a great extent depending on the components of food consumed, residue levels, procedures involved in the preparation of the food, therefore typical values cannot be given. The ranges of relative uncertainties of the main influencing factors, based on the currently available information, are as follow: amount of food consumed: 30 - 90%; recipes of composite foods: 30 - 140%; processing factors 10 - 300%; sampling of plant materials, assuming minimum sample size specified by the Codex sampling procedure, 20 - 35%; sampling processed solid products ~10%; sampling well-mixed processed liquid products: 0%; sub-sampling of large crops: 7 - 21%; sample processing in optimum case: ~10%; analysis of residues in supervised trials (≤ 15%) and monitoring programmes < 25%.

Keywords: Dietary Exposure Assessment; Uncertainty; Sources and Calculation

Abbreviations


Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues

Introduction

Pesticides are indispensable components of modern intensive agricultural production which should supply sufficient amount and quality of food for the continuously growing population of World on practically constant or decreasing arable lands. Pesticides are also required for protection of harvested crops because without them minimum 30-50% of the stored raw agricultural commodities would be lost. The so called “bio-products” may only provide part of the food for a very small fraction of the population. Pesticides, including insecticides, fungicides, herbicides, grows regulators etc., have very diverse chemical structures, physical-chemical properties and mode of action to control target organisms [1].

The pesticide deposit on treated objects vary at a great extent depending on, for instance, the application technique, positioning of nozzles, spatial arrangements and growth stage of plants, microclimatic conditions during application and the quality of the formulated products [2-7]. In addition, heavy rain or sprinkling irrigation can wash off the residues from the treated surface [8,9]. As a result, the range of residue concentrations in individual crop units or primary sample increments (≤ about100 g) taken from a field may be around two magnitudes [10-12]. Horváth and co-workers analysed the distribution of residues in over 20000 primary crop units representing 182 crop-pesticide combinations. They concluded that the within field variation of normalised residues can be well characterised with a lognormal distribution (mean = μ = 1 and standard deviation = σ = 0.8) [13].

Normalized residues: the individual residues making up one dataset are divided by their average value.

Most of the pesticides are toxic substance, which may adversely affect the human health. Therefore, the active ingredients and formulated products are subject to rigorous toxicological and biological efficacy tests implemented according to the OECD Guidance Documents to assure that studies performed in different institutes and test facilities of manufacturers provide comparable and reliable results [14]. Based on the toxicological end-points the no-effect level and or benchmark dose are determined which provide the basis for establishing, applying appropriate safety factors, the acceptable daily intake (ADI mg residue/kg body weight/day) and, if necessary the acute reference dose (ARfD mg residue/kg bw) [15].

Acceptable daily intake (ADI): The estimate of the amount of a chemical in food or drinking-water, expressed on a body weight basis that can be ingested daily over a lifetime without appreciable health risk to the consumer. It is derived on the basis of all the known facts at the time of the evaluation. The ADI is expressed in milligrams of the chemical per kilogram of body weight (a standard adult person weighs 60 kg). It is applied to food additives, residues of pesticides and residues of veterinary drugs in food [16].

Acute reference dose (ARfD): The estimate of the amount of a substance in food or drinking-water, expressed on a body weight basis that can be ingested in a period of 24h or less without appreciable health risk to the consumer. It is derived on the basis of all the known facts at the time of evaluation. The ARfD is expressed in milligrams of the chemical per kilogram of body weight [16].

Point to note: the established ADI and ARfD values are valid only for that product which was subject to toxicological tests, and they may not be applicable for other pesticides having the same active ingredient but have different composition of the technical grade products and formulation [15] due to altered manufacturing process applied by the so called “me too” companies.

The recommended use patterns (including but not limited to dosage rate, time and frequency of application, and the minimum time which must elapse between last application and harvest (called pre-harvest interval, PHI) are determined by national registration authorities to assure effective protection of crops, but avoiding overdose leading to high residues and contamination of the environment [17]. The use pattern providing efficient protection of crops may vary depending on the climatic conditions and agronomical practice in different geographical areas. Before the registration/authorization of a pesticide product, the distribution and magnitude of residues, following the targeted maximum registered/authorised use patterns, are determined with so-called supervised trials conducted, preferably
according to the relevant OECD Guidance Documents [18-20], under strictly controlled conditions on usually small experimental plots ranging from four trees to a few hundred square meters.

The magnitude of residues from trials to trials varies substantially. Detailed analysis of 25766 residue values derived from 1950 supervised trial datasets consisting of a minimum of five residue values, selected by the experts of FAO/WHO Joint Meeting on Pesticide Residues (JMPR) for estimation of supervised trial median residues (STMR), highest residues (HR) and recommending maximum residue levels, revealed that the typical relative standard deviation of field to field variation of residues can be expected to be about 80% [21]. The number of supervised trials submitted for evaluation by the JMPR is most frequently 6 - 8, ranging from three to over 20. These datasets are only 'samples' taken from the widely varying magnitude of parent population of residues present in/on treated crops following the use of pesticide according to the use pattern resulting in the highest residues level. If another set of trials were conducted quite different residue values could be obtained. The magnitude of residues depends mainly on the dosage and PHI.

Before a product is authorized for use, a preliminary exposure assessment should be carried out based on the results of supervised trials. Deterministic models provide simple exposure modelling tools that rely on fixed values derived from data or other information sources. In the context of dietary exposure assessments, the term 'point estimates' refers to a method whereby a fixed value for food consumption (such as the average or high level consumption value) is multiplied by a fixed value for the residue concentration (often the median residue level, or the highest residue observed in supervised trials or MRL specified in the use authorization) and the intakes from all sources are then summed [22]. Although 'point estimates' can be applied at all levels of assessment, they are commonly used as a first step or screening assessment in dietary exposure assessments to eliminate cases where more sophisticated modelling is not required or where more detailed data are unavailable. They are often viewed as efficient models for regulatory decision making because they are relatively simple, rapid and inexpensive to carry out and default values can be applied against upper percentile values of the substance of interest [23,24].

Deterministic methods are used at international level by the JMPR utilising the GEMS/Food cluster diet [25] dividing the World into 17 regions based on similarities of dietary patterns within each cluster [26]. Though the food consumption database has many uncertainties and limitations, it is the best available source of data for predicting long-term intake. JMPR is not a regulatory body, if the dietary exposure exceeds the toxicological reference values and there is no alternative national use pattern reflected by sufficient number of lower residue levels which would lead to lower intake values, the JMPR indicates the case to the Codex Committee on Pesticide Residues (CCPR). CCPR and the Codex Alimentarius Commission are the international risk manager bodies to make decision on the acceptability of risk associated with the reported exposure level [27].

The Primo model [28] used by EFSA for dietary exposure assessment includes the food consumption data from some European Union (EU) member countries. The pesticide active substance is authorized only if the dietary exposure does not exceed the ADI or ARfD values in any of the countries reported consumption data. On the other hand, all routes of potential aggregate exposure are assessed with a complex probabilistic methodology in USA [17] and the authorization of the use of the given product depends on the outcome of the assessment.

The requirement of increased transparency of risk assessment of food and feed raised the need for characterising the uncertainties and communicate them to risk managers so they can determine when to take appropriate measures [29]. Assessments must say clearly and unambiguously what uncertainties have been identified and what is their impact on the overall assessment outcome.

The term uncertainty has different meanings and definitions [16,30]. In general terms, the uncertainty is defined as all types of limitations in the knowledge available to assessors at the time of assessment and within the time and resources available for the assessment [31]. In metrology, the uncertainty of measurement is defined as a parameter, associated with the results of measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand. The parameter may be a standard deviation or
Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues

a given multiple of it [32]. The scientifically correct interpretation of the quantitative results requires information on the range within which the estimated exposure can be expected in 95% of the cases.

Sources of uncertainties in risk assessment can be grouped by scenario, model and parameter [33] or as outlined by EFSA in 2006 [29] by the “assessment objectives, the exposure scenario(s), the exposure model, the model inputs, and the performance of the assessment”. Types of uncertainties include those coming from the scenario (like processing information, recipe data, food conversion factors, market share and occurrence), from parameters (like input data and their availability, concentration data, food consumption data, conversion/processing factors), or from the model chosen (like extrapolation of short term dietary surveys to chronic exposure [34]).

Due to their conservative nature and use as first tier assessment tools, typically uncertainty analysis will not be undertaken for deterministic models. However, to ensure that such confidence is appropriate, deterministic/point estimates methods should undergo a thorough evaluation of their inherent uncertainties. Furthermore, exposure assessor applying such models should be aware of the uncertainty in the data used in such models and the impact this may have on the initial estimate [35].

The objectives of our work are to account for the uncertainties associated with the deterministic dietary exposure assessment, and show some examples how the missing information can be estimated based on available experimental data and expert knowledge.

Materials and Methods

Deterministic models for calculation of dietary exposure

When the deterministic model is used, the long-term international (IEDI) or national daily intake (NEDI) are calculated from the median residue obtained from supervised trials (STMR) or processing studies (STMR-P) and the relevant average consumption data (F) [36]:

\[ IEDI = \sum (STMR_i \times F_i) \text{ or } \sum (STMR - P_i \times F_i) \]  

The short-term intake, calculated from the food consumption within 24 hours, taking into account the large portion sizes (LP=97.5\textsuperscript{th} percentile consumption of eaters) reported from food consumption surveys, preferentially expressed as consumed food [kg]/body weight[kg].

Calculations of short-term intake recognize four different cases (1, 2a, 2b and 3) [36].

In Case 1, the residue in a composite sample (raw or processed) reflects the residue level in a meal-sized portion of the commodity (unit weight, U, is below 0.025 kg). Case 1 also applies to meat, liver, kidney, edible offal, and eggs, and for grains, oil seed, and pulse commodities when the estimates are based on post-harvest use of the pesticide [37].

\[ IESTI = \frac{LP \times (HRorHR - P)}{bw} \]  

where HR is the highest residue detected in supervised trial samples or in processed products (HR-P), bw is the body weight of the eater of the large portion (LP). Each large portion is derived from a national food consumption survey rather than from cluster diets.

In Case 2, the meal-sized portion of a fruit or vegetable may be higher than its unit weight, U, of 0.025 kg. In this case the consumer would eat more than one crop unit. Extensive research on the variability of residues had shown that the residue level in a unit of fruit or vegetable, e.g., a single apple or a single carrot, may be substantially higher than the average residue in a composite sample taken from a

Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues

lot. This issue was accounted for by the introduction of a variability factor \( v = 3 \) \cite{38} into the international risk assessment \cite{36} by the JMPR. The variability factor is defined as:

\[
v = \frac{R_{0.975}}{\bar{R}} \quad (3)
\]

where \( R_{0.975} \) is the 97.5\textsuperscript{th} percentile of residues in the crop units of the sampled lot and \( \bar{R} \) is the average of residues in the lot. The 97.5\textsuperscript{th} percentile of the distribution of residues \cite{39} cannot be measured directly, but calculated from the highest residue obtained from supervised trials (HR) multiplied by the "variability factor" (\( v \)).

In Case 2a, the unit edible weight of raw commodity (\( U_e \)) is above 25g, but less than large portion weight.

\[
IESTI = \frac{U_e \times (HR_{or} HR - P) \times v + (LP - U_e) \times (HR_{or} HR - P)}{bw} \quad (4)
\]

The Case 2a formula assumes that the first unit contains residues at the \( [HR \times v] \) level and the next ones contain residues at the HR level, which represents the residue in the composite sample taken from the same lot.

In Case 2b, unit edible weight of raw commodity, \( U_e \), exceeds large portion weight.

\[
IESTI = \frac{LP \times (HR \ or \ HR - P) \times v}{bw} \quad (5)
\]

In Case 2b there is only one unit (or part of it) consumed (e.g. a slice of watermelon) and it contains residues at the \( [HR \times v] \) level.

In Case 3, equation 6 is used in case of processed commodities where due to bulking or blending the STMR-P represents the likely highest residue. Case 3 also applies to milk, grains, oil seeds, and pulses for which estimates are based on the pre-harvest use of the pesticide.

\[
IESTI = \frac{(LP \times STMR - P)}{bw} \quad (6)
\]

The calculated long-term (EDI) and short-term (ESTI) values expressed in mg/kg body weight are compared to the corresponding ADI and ARfD [mg/kgbw] values, as appropriate.

General rules of propagation of random errors

Equations 1-6 describe the deterministic model for the calculation of dietary exposure of consumers to pesticide residues. Each parameter included in the equations has its own uncertainty. The overall uncertainties of the calculated long-term and short-term intakes can be estimated from the quantifiable uncertainties of the relevant parameters applying the general rules of propagation of error \cite{40}.

The description of the two basic situations follows:

(a) The result is the sum of separately measured values (\( P, Q, R \)), such as the pesticide residues [mg] in individual food items consumed within a day:

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]

\[ \]
Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues

\[ Y = C_1P \pm C_2Q \pm C_3R \ldots \quad (7) \]

Applying their standard deviations \( (SP, SQ, SR) \) and multiplying factors \( (C_1, C_2, C_3) \), the combined uncertainty of the \( Y \) is calculated as:

\[ S_{Y(s_{p,q,r})} = \sqrt{(C_1 \times s_p)^2 + (C_2 \times s_q)^2 + (C_3 \times s_r)^2} \ldots \quad (8) \]

(b) The result is obtained with multiplication or division:

\[ Y = \frac{K \times P}{Q \times R} \quad (9) \]

The relative standard deviation (coefficient of variation) of the \( P \) value is:

\[ CV_p = \frac{s_p}{P} \quad (10) \]

The relative uncertainty (random error) of the calculated \( Y \) value is calculated as:

\[ CV_y = \sqrt{(k \times CV_p)^2 + CV_Q^2 + CV_R^2} \quad (11) \]

A typical example is the calculation of the uncertainty of measured residue concentration in a sample considering the uncertainty of sampling, sub-sampling, withdrawing the test portion from the comminuted laboratory (sub-)sample and analysis of the residues in test portions. The analytical phase can be further subdivided. Let's assume that the average residue in the laboratory sample is 1 mg/kg. Due to random errors the subsample and test portion contain 1.2 mg/kg and 0.95 mg/kg residue respectively (these concentrations are not known in practice) and the average recovery is 70%. The actually reported concentration, after adjusting the measured residue with the average recovery, will be:

\[ 1 \times 1.2 \times 0.95 / 0.7 = 1.62 \text{ mg/kg}. \]

The combined uncertainty of the reported value can be calculated with equation 11. It should be noted that laboratories usually report the uncertainty of their measurement only based on the results of recovery tests which grossly underestimate the real uncertainty as it will be shown hereunder.

**Results and Discussion**

The uncertainties of input parameters of the deterministic exposure assessment models are characterized based on the available information.

**Food intake (\( F_i \)) and \( LP_i \)**

The uncertainty of the reported consumed portion is affected by several factors such as:
• the applied food consumption collection method (2-day dietary recall, dietary records, etc.);

• representativeness of the sampling design which should cover, for instance, regional, seasonal variations and the specific eating habits of minorities.

The above uncertainties cannot be quantified, but can be reduced with carefully designed stratified random sampling design and collecting sufficient number of samples from each stratum. The number of samples taken from specified strata of the populations (e.g. infants, toddlers, children, adolescents, adults, elderly people, women and men) will determine the validity of the large portions reported. The minimum number of samples which should be taken from each stratum to obtain reliable estimate of the LP (\(\beta_p = 97.5\text{th percentile of eaters}\)) is 119, which provides \(\beta_p = 95\%\) probability for finding at least one value above the P97.5\text{th percentile of the quantity of food consumed by the eaters}\ [41]. Naturally, when drawing random samples, the study organizer would not know how many eaters of a given food would be included. The probability of finding the large portion can only be calculated after the dietary intake interviews have been conducted and evaluated. The relationship of the probability (\(\beta_i\)) of finding at least one value above the selected percentile, \(\beta_p\) and the number of samples is described by equation 12:

\[
\beta_i = 1 - \beta_p^n \text{ or } n = \frac{\lg (1 - \beta_i)}{\lg \beta_p}
\]  

(12)

The probability of obtaining the accurate information for the LP based on the number of eaters interviewed (\(n\)) can be easily calculated with equation 12.

Another source of uncertainty of the estimated LP, or \(F_i\) is the memory of the interviewee enabling to remember what foods were eaten on the previous day(s) and the ability to estimate the portion eaten based on the pictures shown in the picture book and auxiliary aids (household measures, abstract shapes, etc.) used during the interview. The applicability of picture series and portion size measurement aids was tested in so-called validation studies including recall of food consumed during the previous 1 to 4 days. The results, indicating the combined effect of the ability of perception and memory of the interviewed persons, have been published [42-45]. They indicate that the relative uncertainty of estimation of the mass of portion eaten varies between 30% and 90% depending on the type of food and its presentation on the plate compared to that shown on relevant series of pictures on the picture book. The non-quantifiable uncertainty is further increased by two practical limitations:

• limited number of representative food portions can only be depicted in picture book and they should be used for helping the interviewee to estimate the portions of other foods of similar appearance;

• at home, different kind of foods are served in one plate as shown in Figure 1, while various sizes of portions of one kind of food are shown in the picture book (Figure 2) to assist the interviewee to estimate the food consumed.

Figure 1: Noodles and stew made of pork served according to normal practice at home.

Pesticide residue concentrations

The uncertainty of pesticide residue measurement results affects the STMR, HR values obtained from supervised trials and the residues measures in samples collected as part of the monitoring or targeted selective survey programmes. The common features of the determination of pesticide residues and the specific aspects related to residues in various types of samples are discussed separately in the following sections.

The combined uncertainty of the determination of pesticide residues (CV_R) includes the relative uncertainty of sampling (CV_S), reducing the sample mass of large crops (sub-sampling CV_{SS} e.g. taking representative segments from each of the 5 - 10 melon fruits making up the composite sample), sample processing by comminution of sub-sample and removing representative portions for extraction (CV_{sp}), and extraction, clean-up, identification and quantitative determination of residues (CV_A). The CV_R can be calculated with equation 11. 

**Sampling**

The uncertainty of sampling was studied from different databases [10,50,51,52] deriving at approximately similar sampling uncertainty values.

The earlier results were confirmed by Farkas and co-workers [53] who evaluated the results of over 12000 replicate sample sets (25867 individual residue data) derived from supervised trials. The comprehensive analyses of available data enabled the estimation of uncertainty of sampling for 106 individual crops and 24 crop groups according to the Codex Classification of Foods and Feeds [54] together with their upper 95% confidence limits. Further on, it was confirmed that the estimated uncertainties and their confidence intervals are independent from the residue distribution in primary samples and crop units [55] within a single plot or lot. The typical sampling uncertainties of residues in primary samples CV_{S1} (upper confidence limit is given in brackets) ranged from 0.57 (0.72) for tree nuts to 1.6 (2.6) for seeds for beverages and sweets. In the majority of medium and large crops it was around 0.8. The expectable sampling uncertainty of average residues in composite samples including k primary units can be calculated based on the central limit theorem:

\[ CV_{sk} = \frac{CV_{S1}}{\sqrt{k}} \] (13)

**Sub-sampling**

The uncertainty of sub-sampling was studied with post-harvest treated papaya and field treated jackfruits and cucumbers [56]. The concentration of benomyl residues in opposite concentric segments of papaya were between 0.53 and 1.57 mg/kg, while the iprodione and primiphos-methyl residues ranged from to 0.16 to 2.46 mg/kg in opposite segments of cucumber samples. The sample size reduction resulted in relative uncertainties of 17 and 21 % for field treated jackfruits and cucumber and 7% for post-harvest treated papaya indicating that the post-harvest treatment provided more uniform distribution of residues in/on treated fruits.

**Sample processing**

Sample processing aims to provide well-mixed matrix from the laboratory sample enabling to withdraw test portions of a few grams for extraction and further steps of qualitative and quantitative analyses. The uncertainty of sample processing (CV_{sp}) estimated from the residue content of test portions, depends on the mass of test portion, nature of the sample material, maturity of the fruits and vegetables, the equipment used for comminution of samples, the temperature of processing (processing in deep-frozen conditions in the presence of dry ice or liquid nitrogen provides much better homogeneity), but it is independent from the pesticide provided that the residues remain stable during processing. The results of the studies [57-60] show wide variation.

Under the conditions of regular pesticide residue analyses a relative sample processing uncertainty of 10% can be realistically achieved. To try to reduce it would involve substantial time and cost but would not improve the combined uncertainty of the results.

Points to note: The target 10% can be achieved relatively easily when ≥15 g test portions are withdrawn. However, keeping the same sample processing procedure and decrease the test portion size from 15 g to 1 g would increase the sample processing uncertainty 3.9 times according to the basic equation of fundamental sampling error introduced by Gy [61] and applied by Minkkinen [62]:

---

**Citation:** Árpád Ambrus and Júlia Szenczi-Cseh. "Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues". *EC Nutrition* 7.5 (2017): 228-251.
Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues

$$CV_{sp}^2 = Cd^3 \left( \frac{1}{M_{tp}} - \frac{1}{M_L} \right)$$  \hspace{1cm} (14)

where C is the shape factor, d is the upper 95% of the particle size in the comminuted matrix, $M_{tp}$ is the mass of test portion and $M_L$ is the mass of comminuted matrix.

The targeted $CV_{sp}$ of 10 % may not be applicable for difficult sample materials with hard peel and soft pulp, such as tomato. Their processing requires special attention.

Analysis of test portions is carried out with many different methods with varying recovery. The review of recovery values reported by the JMPR for various compounds indicated that on an average the uncertainty of analyses can be assumed to be ≤ 15%.

In summary, considering the typical uncertainties of determination of pesticide residues in composite samples derived from supervised trials are sampling $CV_s = 0.25$, sub-sampling of large crops $CV_{ss} = 0.2$, sample preparation $CV_{sp} = 0.1$ and analysis $CV_{A} = 0.15$, the combined uncertainty of determination of residues in large crops:

$$CV_a = (0.25^2 + 0.2^2 + 0.1^2 + 0.15^2)^{1/2} = 0.37$$ \hspace{1cm} (15)

and, similarly, for medium and small crops 0.31.

Supervised trial median residue value (STMR)

The results of supervised trials, conducted according to that national use pattern (called critical Good Agricultural Practice, GAP, in the JMPR reports), which leads to the highest residue levels in/on the treated crops, are selected for estimation of the supervised trial median residue (STMR), the highest residue (HR) and the maximum residue level (MRL).

There are two different sources of uncertainties affecting the reliability of reported STMR value:

- determination of pesticide residues;
- number of supervised trials.

Regarding the number of supervised trials, the data sets representing the residues deriving from the use of pesticide according to the specified use patterns is only one sample drawn from the population of residues. The uncertainty of the median residue depends on the number residue values making up the dataset. The approximate 95% range of median values are summarised in Table 1, which shows the rank number of ordered residue values encompassing the 95% range with 95% probability [63].

<table>
<thead>
<tr>
<th>Number of samples</th>
<th>Rank # of lower boundary</th>
<th>Rank # of upper boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>25</td>
<td>7</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 1: Rank number of residue values encompassing the 95% range of median value.

Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues

The table indicates that in case of 6 to 8 trials, which most frequently make up the supervised trial datasets, the true median value can be anywhere between 0 and the maximum residue in the dataset, and for 10 trials the median can be at or below the 9th highest residue value in 95% of the cases. The probability of that the median corresponds with one of the values within the dataset is not equal as shown in Figure 3, which depicts the probability distribution of median residue, using normal approximation, in case of 18 supervised trials conducted in USA where bifenthrin was applied at 0.22 kg active substance/hectare on strawberry fields [64].

The approximate standard deviation of the selected percentile of the residue data population obtained from supervised trials can be calculated with the general equation of standard deviation of binominal distribution as:

$$S_p = \sqrt{N \times p \times q} \quad (16)$$

where $p$ is the selected percentile, $q = 1 - p$ and $N$ is the number of data points. In case of median the $p = q = 0.5$. The relationship is accurate for $N \geq 20$ values, but provides approximate value for smaller $N$ values [63]. The approximate relative uncertainty of the STMR value can be calculated, assuming normal distribution, from the 95% range of residues ($R_{P0.975} - R_{P0.025}$) in the dataset divided by the median value ($S_{STMR}$) as:

$$S_{STMR} = \frac{R_{P0.975} - R_{P0.025}}{2 \times 1.96} \quad (17)$$

$$CV_{STMR} = \frac{S_{STMR}}{STMR} \quad (18)$$

The combined relative uncertainty ($CV_{comb}$) of the STMR value is calculated from the $CV_{STMR}$ and $CV_R$ (equation 15) with equation 11:

$$CV_{comb} = \sqrt{CV_{R}^2 + CV_{STMR}^2} \quad (19)$$

Figure 3: Probability of the median corresponding to the ranked residue (rank number of highest residue is 1). The 95% confidence intervals of the bifenthrin residues are between 0.36 mg/kg and 0.86 mg/kg (12th and 5th rank numbers).
Highest residue (HR) value

The factors affecting the determination of pesticide residues in a sample containing the highest residue in the dataset are the same as described in the previous section.

The HR is expected to represent the 97.5th percentile of the residue population resulted from the application of a pesticide according to the critical GAP. It cannot be reliably determined from datasets most frequently consisting of 6 - 8 residue values. To estimate the 97.5th percentile, the relationship between the median residues (M) and the highest residues was studied in datasets which contained minimum five residue values and less than 50% of them were below the LOQ [21]. The database included 25766 trials and 1950 pesticide-crop combinations representing a wide range of practical conditions. To be able to compare the spread of residues of largely varying magnitude, the residue values (R) within each dataset were normalised (R x R^-1). The spread of residues was characterised by arranging them in R < M; M ≤ R < 4M; 4M ≤ R < 5M; 5M ≤ R < 6M; 6M ≤ R < 7M and R ≥ 7M intervals. The cumulative frequency of residues in median ranges is summarised in Table 2.

<table>
<thead>
<tr>
<th>Percentage of data sets in the median (M) ranges</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>R &lt; 3M</td>
<td>54.50</td>
</tr>
<tr>
<td>3M ≤ R &lt; 4M</td>
<td>71.61</td>
</tr>
<tr>
<td>4M ≤ R &lt; 5M</td>
<td>78.58</td>
</tr>
<tr>
<td>5M ≤ R &lt; 6M</td>
<td>85.92</td>
</tr>
<tr>
<td>6M ≤ R &lt; 7M</td>
<td>88.68</td>
</tr>
<tr>
<td>7M ≤ R</td>
<td>100.00</td>
</tr>
</tbody>
</table>

*Table 2: Percentage distribution of residues in median ranges.*

The results indicate that about 89% of the residues were < 7M. Consequently, the HR values being within 3-5M range likely underestimate the true 97.5th percentile of the parent population. Horváth and Ambrus [65] studied the relationship between the median residues in datasets and the 97.5th percentile of the parent population of residues in composite samples by drawing 10000 random samples of sizes 4 - 32 with replacement from the normalized parent population of 25766 supervised trial residue data (mean=1; S = 0.974, median = 0.823, P_0.975 = 3.009; max = 9.601). The ratios of the known 97.5th percentile of the parent population and the 5th percentile of medians P_0.05*M in random samples were calculated for each sample size

\[ f_{M,n} = P_{0.975}/P_{0.05*M,n} \]

The relationship of the \( f_{M,n} \) and number of samples in datasets (n) could be described with a second order equation:

\[ f_{M,n} = 10.233n^{-0.228} R^2 = 0.9909 \]

(20)

The HR_{0.975} can be calculated as:

\[ HR_{P0.975} = f_{M,n} \times STMR_{0.975} \]

(21)

The \( f_{M,n} \) values for sample sizes n = 3 to 15 are given in Table 3.

<table>
<thead>
<tr>
<th>n</th>
<th>f_{M,n}P0.975</th>
<th>n</th>
<th>f_{M,n}P0.975</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>8.0</td>
<td>8</td>
<td>6.4</td>
</tr>
<tr>
<td>4</td>
<td>7.5</td>
<td>10</td>
<td>6.1</td>
</tr>
<tr>
<td>5</td>
<td>7.1</td>
<td>12</td>
<td>5.8</td>
</tr>
<tr>
<td>6</td>
<td>6.8</td>
<td>15</td>
<td>5.5</td>
</tr>
</tbody>
</table>

*Table 3: The f_{M,n} values for calculation of expectable HR values with 95% probability.*

---

**Citation:** Árpád Ambrus and Júlia Szenczi-Cseh. "Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues". *EC Nutrition* 7.5 (2017): 228-251.
The results reflect the higher uncertainty of estimation of the 97.5th percentile based on small datasets.

To limit the overestimated values, where the HR observed in a dataset is at or above 6M, we can assume that the observed value properly represents the likely 97.5th percentile of the parent population and it can be directly used for the calculation of short-term intake. It is recognised that the HR_{90,975} values calculated with equation 21 would still overestimate the true 97.5th percentile of the parent residue population in about 35% of the cases. However, underestimation of the acute exposure may have severe consequences, therefore overestimation is justified.

Variability factor

The variability factor, \( v \), was determined from residues in/on unit crops taken from lots of marketed fruits and vegetables [11] Hill and specifically designed field trials [65,66,67]. The JMPR 2005 JMPR confirmed the previously estimated variability factor of 3 [68]. The residue distribution in crop units taken from 182 independent lots was analysed [55]. Each sample set consisted of 90-320 residue data amounting to 20999 residue values. The average variability factor calculated from the normalised residues in each dataset was 3.07. The standard deviation (expressed in rank numbers) of the 97.5th percentile of residues calculated with equation 16 was 22.62. The corresponding 95% confidence intervals of the variability factor are 2.994 and 3.166. The relative standard deviation of the estimated variability factor is:

\[
CV_v = \frac{(3.166 - 2.994)}{2 \times 1.96 \times 3.0745} = 0.0141
\]  

(21)

The probability distribution of estimated P0.975 percentile of 20999 normalised residues is shown in Figure 4.

---

**Figure 4:** Probability distribution of the estimated P0.975 percentile of normalised residues.
Use of monitoring data for estimation of dietary intake

The primary purpose of conducting monitoring programmes is to verify that the pesticides are used according to the corresponding authorized use patterns, and the residues do not exceed the registered maximum residue limit (MRL mg/kg). For this purpose, the residues defined for enforcement purposes are looked for and determined. The measured residues can be directly used for estimation of dietary intake only if the whole fruit or vegetable is consumed and the residue definition for enforcement and dietary risk assessment purposes is the same.

Residues measured in edible portion

The simplest situation is when the residues defined for dietary intake assessment are also measured in the edible portion of the commodity during the supervised trials. In this case, the uncertainty of the measured residue can be calculated as described for pesticide residues in general. Special attention is required for cases where the residue components are determined separately, because the reproducibility standard deviations of each residue components must be combined according to equation 8, and the \( CV_A \) should be calculated from the combined standard deviation and the sum of measured residues. Then the \( CV_A \) incorporating the uncertainties of the analyses of all residue components can be used for calculation of \( CV_R \) with equation 11 as illustrated with equation 15.

Residue definitions for enforcement and dietary intake calculation are different

The example of fluxapyroxad residues evaluated by the 2015 JMPR [69] illustrates the complexity of such cases. The definition for residues for testing compliance with MRLs is fluxapyroxad, but the definition of residues for dietary intake calculations is the sum of fluxapyroxad, 3-(difluoromethyl)-N-(3’4’,5’-trifluoro-1,1’-biphenyl-2-yl)-1H-pyrazole-4-carboxamide (M700F008), and 3-(difluoromethyl)-1-(β-D-glucopyranosyl)-N-(3’4’,5’-trifluoro-1,1’-biphenyl-2-yl)-1H-pyrazole-4-carboxamide (M700F048), expressed as fluxapyroxad for plant commodities (Figure 5).

The proportion of parent compound and its metabolites (M700F008 and M700F048) varies in treated crops after the treatment. No correlation could be found between them and the total residue as shown in Figure 6. The total residue should be used for the calculation of long- and short-term dietary intakes. If the intake calculations using the residues of parent compound determined in treated commodities approaches the ADI or ARfD values, selective field surveys, taking samples from fields known to be treated with the pesticide, should be initiated and the residues included in the residue definition should be determined in the edible portions of treated crops [65].

**Figure 5:** Chemical structure of parent fluxapyroxad and its major plant metabolites.

In some cases, residues defined for enforcement and dietary exposure assessment purposes are interrelated. For instance, the JMPR defined [69] the residue for enforcement in plant commodities as spirotetramat plus spirotetramat enol, expressed as spirotetramat; for dietary intake calculation: spirotetramat plus the metabolites enol, ketohydroxy, enol glucoside, and monohydroxy, expressed as spirotetramat (Figure 7).

**Figure 6:** Ratio of total residue and metabolite as a function of total residue/parent compound in cherry 1-3 days after application of fluxapyroxad.

**Figure 7:** Spirotetramat and metabolites included in the residue definition for dietary intake calculations in plants JMPR 2008.
For instance, the residues were measured in sweet corn according to both residue definitions. Good linear correlation was found [65] between the corresponding concentrations that can be described as:

$$C_{\text{risk}} = 1.23 \times C_{\text{enforc}} + 0.041; R^2 = 0.9785 \quad (22)$$

In such cases, the residue measured in samples taken for enforcement purposes can be used to calculate the corresponding residue concentration for calculation of dietary intake with equation 22.

**No residue data are available in edible portion**

There are several crops eaten as prepared for residue analysis. For instance, berry fruits, pome fruits, lettuce, etc. In other cases, (e.g., watermelon, kiwi fruit, stone fruits) only a portion of the commodity is consumed. Extrapolation of residues from the whole commodity to its edible portion may be possible, based on the results of metabolism studies, in some cases, especially when all residues remain on the surface. In other cases, the intake calculation should be carried out with the residues measured in the whole commodity (e.g., grapefruit versus its edible pulp), which may lead to overestimation of the intake with unquantifiable uncertainty (lack of knowledge and information), because the residues in whole crop are usually higher than in the edible portions. In critical cases, selective surveys should be initiated as it was suggested above for cases where the residue definitions are different.

**Effect of processing**

Many raw agricultural products undergo different kinds of processing (e.g., milling, backing cooking etc.) before they are used for human consumption. During processing the pesticide residues may be evaporated, concentrated and degraded due to chemical and microbiological reactions. Therefore, taking into account the effect of processing in dietary risk assessment is very important.

Studies on the effects of processing on the magnitude and composition of pesticide residues revealed that the main reactions are hydrolytic, because processes involving heating would generally inactivate enzymes present in the commodity. Since the substrate, itself is not likely to have a major effect, the presence of the commodity during such studies, is not required. The OECD Guidance document on studying the effect of processing [70] summarised the typical conditions for performing the nature and magnitude of residues studies to simulate the effect of processing (Table 4).

<table>
<thead>
<tr>
<th>Type of process</th>
<th>Critical operation</th>
<th>Temperature (°C)</th>
<th>Time (min)</th>
<th>pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooking vegetables, cereals</td>
<td>Boiling</td>
<td>100</td>
<td>15 - 50</td>
<td>4.5 - 7</td>
</tr>
<tr>
<td>Fruit preserves</td>
<td>Pasteurisation</td>
<td>90 - 95</td>
<td>1 - 20</td>
<td>3 - 4.5</td>
</tr>
<tr>
<td>Vegetable preserves</td>
<td>Sterilisation</td>
<td>118 - 125</td>
<td>5 - 20</td>
<td>4.5 - 7</td>
</tr>
<tr>
<td>Fruit Juice</td>
<td>Pasteurisation</td>
<td>82 - 90</td>
<td>1 - 2</td>
<td>3 - 4.5</td>
</tr>
<tr>
<td>Oil</td>
<td>Raffination</td>
<td>190 - 270</td>
<td>20 - 360</td>
<td>6 - 7</td>
</tr>
</tbody>
</table>

**Table 4: Typical conditions of food processing.**

Processing studies aim to represent the most important industrial and household practices and operations, such as the manufacture of vegetable oils, milling cereal products and preparing bread; the production of juices, wine; washing, peeling, blanching, boiling, blending and deep-frying, etc. The effect of processing is described by the so called processing factor defined as:

$$P_f = \frac{\text{concentration of residues in processed commodity} \left[ \frac{mg}{kg} \right]}{\text{concentration of residues in raw agricultural commodity} \left[ \frac{mg}{kg} \right]} \quad (23)$$

In some cases, the processing factor is indicated with less than sign "<". It indicates that the residues were present below the limit of quantification (LOQ) in the processed product. Processing factors are very much affected by the characteristics of pesticide residues such as water or fat solubility, the distribution of the pesticide on the commodity, e.g., surface or systemic, or its application in pre- or post-harvest treatments are also relevant. Therefore, the processing factor should be considered as a combination of the process, pesticide residue and the commodity [37].

The Pf depends on the processing yield. It should be distinguished from the yield factor which describes the mass ratio of processed product (e.g. flour) and the initial, usually, raw commodity (e.g. wheat grain).

\[
Y_f = \frac{\text{Mass of processed commodity [kg]}}{\text{Mass of raw agricultural commodity [kg]}}
\]  

(24)

Processing can lead to an increase or a decrease in residues, depending on the specific processing conditions and physicochemical properties of the active substance.

The results of processing studies are influenced by many parameters including the processing method, maturity and variety of raw product, the yield of the process (e.g. the quality and quantity of flour obtained by milling), the oil content of seeds, etc. Consequently, when several studies are conducted their results can widely vary. The reports of processing studies are submitted to the national registration authorities to support the claim for the authorization of their pesticide products. The reports and the results of the studies are part of the confidential data package and not publicly available. The FAO/WHO JMPR publishes the summary of processing studies, the processing factors and the estimated STMR-P values in processed products. To fill the gap, the German Federal Institute for Risk Assessment (BfR) prepared a comprehensive database providing more detailed information related to over 6500 processing factors for a total of 190 pesticide active ingredients based on the studies validated by its experts [71]. The database also includes information on the effect of washing cleaning and partition of residues between peel and pulp of fruits.

Though the German database contains large number of processing factors, it may be the case that it does not contain information for a particular pesticide – commodity – process combination. The OECD Guidance document [72] provides guidance for the possible extrapolation of the Pf from one commodity to other based on 19 processing procedure. In addition to the processing procedure, the distribution of pesticide residues in/on treated crops, water solubility and the octanol water partition coefficient, usually expressed in logarithmic scale logPow, shall be taken into account when results of processing studies are applied for other combinations of pesticide-crop-process. For compounds having very low water solubility (< 0.01 mg/L) the hydrolytic degradation is unlikely. The logPow provides information on fat solubility of the compound. If the logPow is larger than 3, it is likely that the residue will not be transferred to the liquid phase and mostly concentrate in the remaining part of the processed commodity (e.g apple juice ↔ apple pomace).

The extrapolation of the effect of processing from one crop to another must be made with care considering all available information. For instance, Table 5 shows some examples for the effect of washing of fruits. The washing can remove major part of water soluble surface residues indicated by low processing factors (captan, mandipropamid), while washing is practically ineffective if the residues of systemic active substances are in the fruit or they have very low water solubility even if the majority of the residues remain on the surface of the fruits (see lower part of Table 5).
Table 5: Effect of washing on the residues on raw commodity (home procedures).

The octanol-water partition coefficient has a more pronounced influence on the distribution of residues between fruit juice and pomace, as shown in Table 6. In case of compounds with logPow larger than 4 the residues will concentrate in the pomace.

Table 6: Examples for the distribution of residues between fruit juice and pomace.
Because of several factors affect the outcome of processing, the processing factors usually show wide variation as can be seen in Tables 5 and 6. In case of valid study conditions, the occurrence of the processing factors observed for a given pesticide commodity combination have equal probability. Therefore, their standard deviation, $S_{pf}$, is calculated, assuming rectangular distribution [40], from the difference between the maximum and minimum $P_f$ values as:

$$S_{pf} = \frac{P_{f \text{ max}} - P_{f \text{ min}}}{2 \times \sqrt{3}}$$  \hspace{1cm} (25)

Each set of processing studies represents a sample of the unknown population of processing factors. The best estimate of their uncertainty, expressed as relative standard deviation is the pooled variances of the relevant sets of processing factors. However, the magnitudes of median processing factors, $M_{pf}$, are different, therefore the calculation shall be carried out with the relative standard deviations $CVP_{pf}$ using the median processing factor as a robust estimate of the mean value. The advantage of using the median value is that it is not affected by the extreme values:

$$CVP_{pf} = \frac{S_{pf}}{M_{pf}}$$  \hspace{1cm} (26)

$$CVP_{pf \text{ pooled}} = \sqrt{\frac{\sum_{i=1}^{k} df_i \times CVP_{pf i}^2}{\sum_{i=1}^{k} df_i}}$$  \hspace{1cm} (27)

Where $k$ is the number of datasets, $CVP_{pf}$ is the calculated relative standard deviation of the $i$th dataset, $df_i$ is the degree of freedom of the $i$th dataset. Part of the calculation is shown in Table 7.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Processing factor</th>
<th>No</th>
<th>$S_{pf}$</th>
<th>$CVP_{pf}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Median</td>
<td></td>
</tr>
<tr>
<td>Captan²</td>
<td>0.27</td>
<td>1.05</td>
<td>0.68</td>
<td>9</td>
</tr>
<tr>
<td>Captan³</td>
<td>0.02</td>
<td>0.83</td>
<td>&lt; 0.50</td>
<td>9</td>
</tr>
<tr>
<td>Novaluron</td>
<td>0.003</td>
<td>0.091</td>
<td>&lt; 0.010</td>
<td>6</td>
</tr>
<tr>
<td>Trifloxystrobin</td>
<td>0.06</td>
<td>0.29</td>
<td>0.10</td>
<td>6</td>
</tr>
<tr>
<td>Tebufenpyrad</td>
<td>0.63</td>
<td>1</td>
<td>&lt; 0.71</td>
<td>6</td>
</tr>
</tbody>
</table>

**Table 7**: Calculation of typical relative uncertainty of processing apples to juice.

Notes: 1. Rounded values are presented
2. Juice is unclarified
3. Juice is unclarified and pasteurized
4. Atypical value due to very low median $P_f$ of < 0.01

The novaluron median processing factor is < 0.01 and makes the calculated relative standard deviation atypical. Therefore, it has to be left out from the calculation of the pooled relative standard deviation with equation 27 giving $CVP_{pf \text{ pooled}} = 0.399$. Since the processing methods used are not known in dietary surveys, the pooled $CVP_{pf}$ can be used, as the best estimate in the calculation of the uncertainty of

Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues

processing factor, which is affected by the raw and processed product and the processing operations. Care should be taken to select only the relevant sets of processing studies when the pooled CV_p is calculated.

In some cases, the definition of residue may include metabolites of different water solubility and P_w values. In such cases, the extrapolation of results obtained in one commodity to another would require especial attention and careful evaluation of all available information. Otherwise erroneous results could be obtained.

Conclusions

The calculation of dietary exposure of consumers to pesticide residues with deterministic methods is relatively simple task if appropriate data are available. However, the estimation of the uncertainty of the point estimate given by the deterministic procedure is very complex task and requires expert judgements in many instances. The detailed analysis of the factors, which may affect the uncertainty of the calculated exposure, underlines the importance of a thorough evaluation of their inherent uncertainties. Furthermore, exposure assessor applying such models should be aware of the uncertainty in the data used in such models and the impact this may have on the initial estimate [35].

The uncertainties of parameters influencing the calculated dietary exposure vary at a great extent depending on the components of food consumed, residue levels, procedures involved in the preparation of the food, therefore typical values cannot be given. The ranges of relative uncertainties of the main influencing factors, based on the currently available information, are as follow: amount of food consumed: 30-90%; recipes of composite foods: 30 - 140%; processing factors 10 - 300%; sampling of plant materials, assuming minimum sample size specified by the Codex sampling procedure [41], 20 - 35%; sampling processed solid products ~10%; sampling processed well-mixed liquid products: 0%; sub-sampling of large crops: 7 - 21%; sample processing in optimum case: ~10%; analysis of residues in supervised trials (≤ 15%) and monitoring programmes < 25%.

Conflict of Interest

The authors report no conflicts of interest.

Bibliography


Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues


Principles of Estimation of Combined Uncertainty of Dietary Exposure to Pesticide Residues


71. German Federal Institute for Risk Assessment (BfR) Data Collection on Processing Factors.