European Food Safety Authority

EFSA Journal 2012;10(10):2839

SCIENTIFIC OPINION

Guidance on the Use of Probabilistic Methodology for Modelling Dietary Exposure to Pesticide Residues[1]

EFSA Panel on Plant Protection Products and their Residues (PPR)[2, 3]

European Food Safety Authority (EFSA), Parma, Italy

SUMMARY

The European Food Safety Authority (EFSA) asked the Panel on Plant Protection Products and their Residues to provide guidance on methodology for performing probabilistic dietary exposure assessment of single or multiple active substances, as a potential additional tool to supplement or complement the standard deterministic methodologies which are currently used in the EU for conducting dietary exposure assessments for pesticides.

Specific guidance is provided for basic assessments but not for refined assessments, where specialised expertise is required to select methods appropriate to the assessment in hand.

The guidance includes probabilistic methods for quantifying some of the major sources of variability and uncertainty affecting dietary exposure to pesticides. Other important sources of variability and uncertainty might be quantified probabilistically in refined assessments but are addressed more simply in basic assessments by conducting alternative model runs with optimistic and pessimistic assumptions.

Guidance is provided on problem formulation, including definition of appropriate scenarios for acute and chronic exposure assessment in the differing contexts of approval of new substances, MRL setting, authorisation of products, evaluation of residues found above the MRL, and annual reviews of residue monitoring data.

[1] On request from EFSA, Question No EFSA-Q-2008-754, adopted on 21 June 2012.

[2] Panel members: Jos Boesten, Claudia Bolognesi, Theo Brock, Ettore Capri, Anthony Hardy, Andy Hart, Karen Ildico Hirsch-Ernst, Susanne Hougaard Bennekou, Robert Luttik, Michael Klein, Kyriaki Machera, Bernadette Ossendorp, Annette Petersen, Yolanda Pico, Andreas Schäffer, Paulo Sousa, Walter Steurbaut, Anita Strömberg, Maria Tasheva, Ton Van Der Linden And Christiane Vleminckx. One member of the Panel did not participate in part of the discussion on the subject referred to above because of potential conflicts of interest identified in accordance with the EFSA policy on declarations of interests.
Correspondence: pesticides.ppr@efsa.europa.eu

[3] Acknowledgement: The Panel wishes to thank the members of the Working Group on Probabilistic Methodology Mandate for the preparation of this opinion: Andy Hart, Bernadette Ossendorp, Paul Hamey, Jacob van Klaveren, Marc Kennedy, David Miller, Annette Petersen, Yolanda Pico, Anita Stromberg, Hilko van der Voet and EFSA staff member Mr. Luc Mohimont for the support provided to this scientific opinion.

Suggested citation: EFSA Panel on Plant Protection Products and their Residues (PPR); Guidance on the Use of Probabilistic Methodology for Modelling Dietary Exposure to Pesticide Residues. EFSA Journal 2012;10(10):2839. [95 pp.] doi:10.2903/j.efsa.2012.2839. Available online: www.efsa.europa.eu/efsajournal

Detailed guidance is provided on methodology for probabilistic modelling of acute and chronic exposures and for quantifying variability and uncertainty in food consumption and residues. For basic probabilistic assessments, optimistic and pessimistic assumptions are used for the effects of processing and for residues below the level of reporting.

A separate section is devoted to additional approaches required for modelling exposure to multiple substances (cumulative assessment): the use of relative potency factors to cumulate exposures to different substances, and a basic methodology for addressing gaps in data on the co-occurrence of residues of different substances.

Specific guidance is provided on the types and formats of outputs that should be produced from a probabilistic assessment. Particular emphasis is placed on characterising the upper tail of the exposure distribution and on 'drill down' techniques to evaluate the reliability of the estimates.

A general approach is recommended for evaluating uncertainties affecting the model outputs. An appendix to the guidance describes uncertainties associated with the methodology recommended in this guidance, and provides a general evaluation of their potential impacts on estimated exposures.

The guidance also includes a checklist of key issues to be considered when writing or peer-reviewing reports on probabilistic exposure assessments, a discussion of approaches to validating probabilistic assessment approaches, and a list of desirable characteristics of software for probabilistic exposure modelling. Some comments are provided on the interpretation of results, while recognising that risk management is outside the remit of EFSA.

Case studies are included in an appendix, illustrating some but not all of the recommended approaches.

Further work will be required to make the methods in this guidance available to end-users in more practical form, including software and more specific user instructions. Some recommendations on this are provided.

KEY WORDS

Probabilistic modelling, dietary exposure assessment, pesticide residues, MRL, monitoring, enforcement, consumer safety, cumulative exposure assessment

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY EFSA

The assessment of dietary exposure to pesticide residues is a key step in process for authorisation of plant protection products and establishment of related maximum residue levels (MRLs) in plant commodities. This is required by Council Directive 91/414/EEC of 15 July 1991 concerning the placing on the market of plant protection products[4], as well as by Regulation 396/2005 of the European Parliament and of the Council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin[5].

Currently, deterministic methods based on WHO guidelines[6,7] are used for assessing dietary exposure. These methods provide an estimation of the exposure of one single virtual consumer and have the advantage of being of simple and fast to use.

In recent years, there has been growing interest internationally in the application of probabilistic techniques to the estimation of exposure to chemicals in food. In contrast with the deterministic methodology, these techniques allow the distribution of intakes[8] amongst multiple individuals in a specified population to be estimated, taking into consideration the variability in food consumption between and within individuals and in occurrence of residues in food commodities.

The European Commission funded research on this methodology from 2000 to 2003 through the Monte Carlo project on the 'Development, validation and application of stochastic modelling of human exposure to food chemicals and nutrients' under the EC Fifth Framework Programme (Quality of Life Key Action 1 on Food Nutrition & Health).

Regarding pesticide residues in particular, the European Commission tendered a project aiming to develop draft guidelines on the use of probabilistic exposure assessment. This resulted in the publication of a report proposing 'guidelines regarding probabilistic exposure assessment in the safety evaluation of pesticides in the EU market[9]'. To date, such guidelines have not been adopted for routine use in decision-making related to authorization of plant protection products or MRL-setting.

The PPR Panel is of the opinion that probabilistic methodology is a potentially useful tool for conducting refined consumer exposure assessments. In particular, in its opinion on cumulative risk assessment[10], the PPR Panel stated that refined cumulative exposure assessments cannot be done without probabilistic methods and recommended that guidance for performing probabilistic exposure assessments should be developed.

TERMS OF REFERENCE AS PROVIDED BY EFSA

The PPR Panel was asked by EFSA to provide guidance on how probabilistic methodologies can be used for estimating dietary exposure, as tools additional to deterministic methods, in the authorisation of plant protection products, in MRL-setting and in the assessment of actual exposure based on residue-monitoring data.

[4] OJ L 230,19.8.1991, now replaced by Regulation (EC) No 1107/2009.
[5] OJ L 70,16.3.2005
[6] WHO/FSF/FOS/97.7: Guidelines for predicting dietary intake of pesticide residues (WHO, 1997a).
[7] WHO/FSF/FOS/97.5: Food consumption and exposure assessment of chemicals (WHO, 1997b).
[8] In the Background provided by EFSA, the word "intake" refers to the amount of chemical taken up by the dietary route, i.e., dietary exposure. In the remainder of this document, "exposure" is used for chemical intake and "consumption" for food intake, to avoid any ambiguity whether "intake" refers to food or chemical.
[9] Boon and Van Klaveren (2003c).
[10] The EFSA Journal (2008) 704, 1-84

ASSESSMENT

1. Introduction

1.1. Interpretation of the Terms of Reference by the PPR Panel

The Terms of Reference provided by EFSA request guidance on probabilistic methodologies for use in the context of authorization, MRL-setting, and assessment of actual dietary exposure of consumers. The specific exposure questions to be assessed in each context differ, and are discussed and defined in section 3.

The Background provided by EFSA includes a reference to cumulative risk assessment. Regulation (EC) No 396/2005 includes a requirement that when suitable methods are available, cumulative exposure from multiple pesticides should also be assessed, as well as exposure to pesticides considered individually. The basic methodology is the same for both types of assessment. Additional methodology specific to cumulative assessments is presented in section 6.

As implied by the terms of reference, the methodologies proposed in this guidance are not intended to replace the existing deterministic methodologies for assessing consumer exposure, but are rather to be seen as complementary, higher tier approaches (see section 2).

1.2. Scope and objectives of the guidance

This guidance proposes a methodology for performing probabilistic dietary exposure assessment of single or multiple active substances in the contexts of authorisation, MRL setting, enforcement actions, and periodic reviews of monitoring data on actual exposures as potential additional tools to supplement or complement the standard deterministic methodologies which are currently used in the EU for conducting dietary exposure assessments. It is designed to provide clear and concise recommendations on key methodological issues that arise in the conduct and review of probabilistic exposure assessments. Further work will be required to make the methods available to end-users in more practical form, including software and more specific user instructions. The Panel understands that EFSA and the Commission will seek to continue this work.

A key feature of the recommended approach is the distinction made between basic and refined probabilistic assessments (see section 2). This document provides specific guidance for basic assessments but not for refined assessments, where it is intended that expert practitioners will select methods appropriate to the assessment in hand. The reasons for this approach are explained in section 2.

The PPR Panel did not consider it appropriate to restrict its recommendations to methodologies already available in ready-to-use software. However, all of its recommendations for basic probabilistic assessments can be implemented without further research and most are available in existing ready-to-use software. Those approaches that are not included in existing software are likely to be added in the near future.

This guidance should support EFSA in performing tasks resulting from Regulations (EC) No 396/2005 and 1107/2009 regarding consumer dietary risk assessments when deterministic approaches are insufficient to reach a risk management decision (see previous section). These methodologies may also be used by governmental bodies and industry in regulatory procedures when considered relevant.

The primary audience for this guidance comprises scientists who need to conduct or evaluate probabilistic exposure assessments at national and EU levels. As such, it is assumed that the reader is

already familiar with types and sources of data on food consumption (e.g EFSA PRIMo 2[11]) and pesticide residues (e.g., EU guidelines 1996/97 Appendix A- I[12]), with basic principles of exposure assessment, and with risk assessment in general. Importantly, it is also assumed that the reader is already familiar with the principles and practices of probabilistic exposure assessment. Introductions to the principles, theory and methods of probabilistic modelling may be found in other publications (e.g., Cullen and Frey, 1999; Vose, 2008; IPCS/WHO, 2008; Van der Voet et al., 2009; Bosgra et al., 2009; Van Klaveren and Boon, 2009; Van der Voet and Slob, 2007; Boon and Van Klaveren, 2003; Pieters et al., 2005; Codex Committee on Pesticide Residues, 2002; U.S. EPA, 1997a). It is also assumed that readers are fully familiar with the technical details of the specific models and software they are using, e.g., from training courses or user manuals.

The document does not address risk management issues such as criteria for acceptable limits to exposure and risk. Deciding what frequency and severity of effects constitute cause for concern and what level of certainty is required (how sure society wants to be) are value judgments that may be influenced by non-scientific considerations and should be made by risk managers not risk assessors. These questions are outside the remit of EFSA, which is confined to risk assessment.

The document is divided into the following main sections:

- Section 2 introduces the Panel's distinction between basic and refined probabilistic assessments.
- Section 3 discusses problem definition, the exposure scenarios to be considered, and the scope of the assessment.
- Section 4 describes the Panel's detailed recommendations for probabilistic modelling of acute exposures to single substances.
- Section 5 describes the Panel's detailed recommendations for probabilistic modelling of chronic exposures to single substances.
- Section 6 describes additional approaches required for modelling exposure to multiple substances (cumulative assessment).
- Section 7 considers the types and formats of outputs that should be produced by a probabilistic assessment.
- Section 8 summarises the recommended approach for evaluation of uncertainties affecting the model outputs.
- Section 9 provides a checklist of key issues to be considered when writing or peer-reviewing reports on probabilistic exposure assessments.

[11] http://www.efsa.europa.eu/en/mrls/mrlteam.htm

[12] EU guidelines
Appendix A- Metabolism in Plants. Commisson of the European Communities 7028/VI/95 rev. 3_22/7 1997
Appendix B- Residue Trials in Plants. Commisson of the European Communities 7029/VI/95 rev. 5_22/7 1997
Appendix C- Rotational Crops. Commisson of the European Communities 7524/VI/95 rev. 2_22/7 1997
Appendix D-Guidelines on comparability, extrapolation, group tolerance and data requirements for setting MRLs, Commisson of the European Communities 7525/VI/95 – rev. 8 ½ 2008.
Appendix E- Processing studies. Commisson of the European Communities 7035/VI/95 rev. 5_22/7 1997
Appendix F- Metabolism in Livestock. Commisson of the European Communities 7030/VI/95 rev. 3_22/7 1997
Appendix G- Livestock Feeding Studies. Commisson of the European Communities 7031/VI/95 rev. 4_22/7 1996
Appendix H- Storage Stability. Commisson of the European Communities 7032/VI/95 rev. 5_22/7 1997
Appendix I-Calculation of MRLs. Commisson of the European Communities 7039/VI/95_22/7 1997

- Section 10 offers some comments on the interpretation of results, without prejudice to risk management judgements which are outside the remit of EFSA.
- Section 11 discusses approaches to validating probabilistic assessment approaches.
- Section 12 summarises desirable characteristics of software for probabilistic exposure modelling.
- The Conclusions section includes some recommendations for further refinement and implementation of the guidance.

Key technical terms used in this guidance are defined in the glossary.

1.3. Case studies

Case studies illustrating most of the approaches recommended for single-substance assessments are presented in Appendix 3. The PPR Panel was not yet able to conduct case studies that follow the proposed approaches in full, because some aspects of the proposed approaches are not yet implemented in ready-to-use software and the Panel lacked the time and resources to program them itself. The purpose of the case studies is to illustrate the general approach, especially the types of outputs and reporting format recommended by the PPR Panel.

2. Tiered approach to probabilistic assessments

Probabilistic approaches are complementary to, and not replacements for, deterministic approaches. They introduce more realism by using distributions to represent the range of variation in consumption, residues, and other relevant parameters rather than using point estimates as in deterministic assessments. They also allow quantification of uncertainties affecting the assessment.

Probabilistic approaches are more complex and require more resources than deterministic approaches. It will therefore be logical to consider them as an option for higher tier assessment in those cases where deterministic approaches are insufficient to reach a risk management decision, e.g., where the deterministic assessment indicates cause for concern and the risk manager wishes to consider more refined estimates to investigate the likelihood and magnitude of exposures above the reference dose. Defining specific criteria for deciding when to conduct probabilistic assessments would require risk management considerations, which are outside the remit of the Panel, but could be an area for further work in consultation with risk managers (see Conclusions).

Rigorous modelling of variability and uncertainty is difficult, requiring refined approaches and advanced statistical expertise to take proper account of the complex nature of variability in the real world, and the many uncertainties that arise from limitations in the types and amounts of data available. This level of analysis is not practical for every assessment. Furthermore, in many cases, basic probabilistic assessments may be sufficient to support a risk management decision. When refined probabilistic assessments are required, they can focus on those sources of uncertainty that have been shown to be important in the basic probabilistic assessment. In both basic and refined assessments, alternative assumptions may be used to explore major sources of uncertainty that remain unquantified. These strategies are explained in more detail below.

2.1. Using optimistic and pessimistic model runs to address uncertainties

When a model component is uncertain, this implies that a range of alternative assumptions could be made for it. Where possible, it would be preferable to represent the uncertainty probabilistically, i.e., as a distribution specifying the probability of each alternative assumption. However, for some uncertainties, specifying probabilities may require refined approaches that are not reasonable to apply in a basic assessment (e.g., specialised statistical modelling and/or the use of expert judgments), and that may not be necessary to reach or support a risk management decision.

A more practical strategy for basic assessment is to carry out alternative model runs using alternative deterministic assumptions for major uncertainties to examine their impact on the estimated dietary exposures. These are referred to here as pessimistic and optimistic model runs.

- **Pessimistic model runs** treat major uncertainties using assumptions that are expected to lead to over-estimation of exposure. The resulting distribution can be considered an upper estimate of the true distribution: this is not an absolute upper bound, but the true exposures are considered unlikely to be higher. If the estimated exposures from pessimistic runs do not exceed the reference dose, then risk managers can be confident that true exposures are unlikely to be of concern. If some of the estimated exposures do exceed the reference dose, then risk managers can be confident that the true proportion of exposures exceeding the reference dose is smaller than the estimated proportion.

- **Optimistic model runs** treat major uncertainties using assumptions that are expected to lead to lower estimates of exposure. For acute assessments, the resulting distribution can be considered a lower estimate of the true distribution: the true exposures are unlikely to be lower. If the estimated exposures deriving from the optimistic runs exceed the reference dose, then risk managers can be confident that true exposures are also likely to exceed the reference dose. If some of the estimated exposures exceed the reference dose, then risk managers can be confident that the true proportion of exposures exceeding the reference dose is larger than the estimated proportion. For chronic assessments the basic optimistic model run is less conservative than the basic pessimistic model, but cannot be guaranteed to under-estimate the true exposure and may be nearly as conservative as the pessimistic model. Nevertheless the optimistic chronic assessment is still useful for indicating when parametric modelling should be considered for the refined assessment (see later, section 5.1.2).

The results of the optimistic and pessimistic model runs can be used to determine whether further refinement of the assessment is useful, as explained in the following section.

It is important to emphasise that both estimates relate to the range of use conditions that are realistically likely to occur.

2.2. Basic and refined probabilistic assessment.

The Panel proposes a tiered approach to probabilistic assessment, as follows:

1. **Basic probabilistic assessment.** The basic assessment comprises two alternative model runs, pessimistic and optimistic, as explained in the preceding section. Sources of variability and uncertainty which are impractical to treat probabilistically in a basic assessment are represented using alternative deterministic assumptions in the pessimistic and optimistic model runs leading, respectively, to upper and lower estimates for the true distribution of exposure. Sources of variability and uncertainty which are practical to treat probabilistically in a basic assessment are represented by the same distributions in both model runs.

   If the results of the pessimistic model raise no concern for risk managers, it can be assumed that the true dietary exposure would also cause no concern, so the assessment can stop[13]. If both the optimistic and pessimistic estimates raise concern and if the level of concern indicates an unacceptable risk, then it can be assumed the true exposure would also raise a similar level of concern. In this case, further refinement is unlikely to be worthwhile if the assessment is acute, whereas in a chronic assessment refinement may require the use of parametric modelling[14]. If the

[13] In this situation, it is not necessary to conduct the optimistic model run.
[14] See sections 2.1 and 5.1.2 for more detail.

pessimistic estimate raises concern but the optimistic estimate does not, it is uncertain whether the true exposure would raise concern, so refined assessment may be helpful[15].

2. **Sensitivity analysis.** Refined assessment will usually involve replacing one or more of the pessimistic elements of the pessimistic basic assessment with more refined assumptions. The choice of which elements to refine may be guided by a simple form of sensitivity analysis: additional models are run with different combinations of the pessimistic and optimistic assumptions from the basic assessment. The purpose of these runs is to help the assessor choose which assumptions to replace with refined modelling in the refined assessment. They should not be used for deciding on the acceptability of the risk because they replace pessimistic assumptions with optimistic assumptions, and are therefore likely to underestimate true exposures.

3. **Refined probabilistic assessment.** Here, pessimistic assumptions of the basic assessment are progressively replaced with refined modelling based on available data and/or expert judgment, taking account of the associated uncertainties[16]. This is likely to require more sophisticated methods than are currently feasible for basic assessment, and specialised expertise. The details are likely to vary case-by-case, depending on the amount and nature of data available and whether extrapolation and/or expert judgment is required. Some pessimistic assumptions from the basic assessment may remain, so the assessments remain somewhat conservative. Optimistic assumptions must not be used in model runs that will be used for deciding on the acceptability of the risk, but could be used for further sensitivity analysis to evaluate the potential value of still further refinements. As the models are progressively refined, the results of the optimistic and pessimistic runs will gradually converge.

The Panel envisages that the basic approaches are suitable for use by anyone who has access to suitable software and is trained in its use. Refined approaches generally involve difficult scientific and statistical issues, and it is recommended that they should be used only within a team possessing expertise in probabilistic modelling and statistics as well as in toxicology, food consumption, pesticide residue behaviour, food preparation and processing, and pesticide usage.

Figure 1 summarises the main steps of the approach recommended by the Panel, including the basic and refined assessments, and optimistic and pessimistic model runs. Refined assessment can be an iterative process, in which different elements of the model are refined progressively until a risk management conclusion is reached. If the assessment indicates cause for concern, options available to risk managers include not only performing refined probabilistic assessment but also collection of further data or risk mitigation.

[15] Other possible options in such cases include collecting additional data to reduce uncertainty, or precautionary management action to reduce the chance of unacceptable exposures.

[16] Although refinements should be designed to improve the realism of the assessment, they will often introduce additional uncertainties, e.g., assumptions regarding the shape of additional distributions.

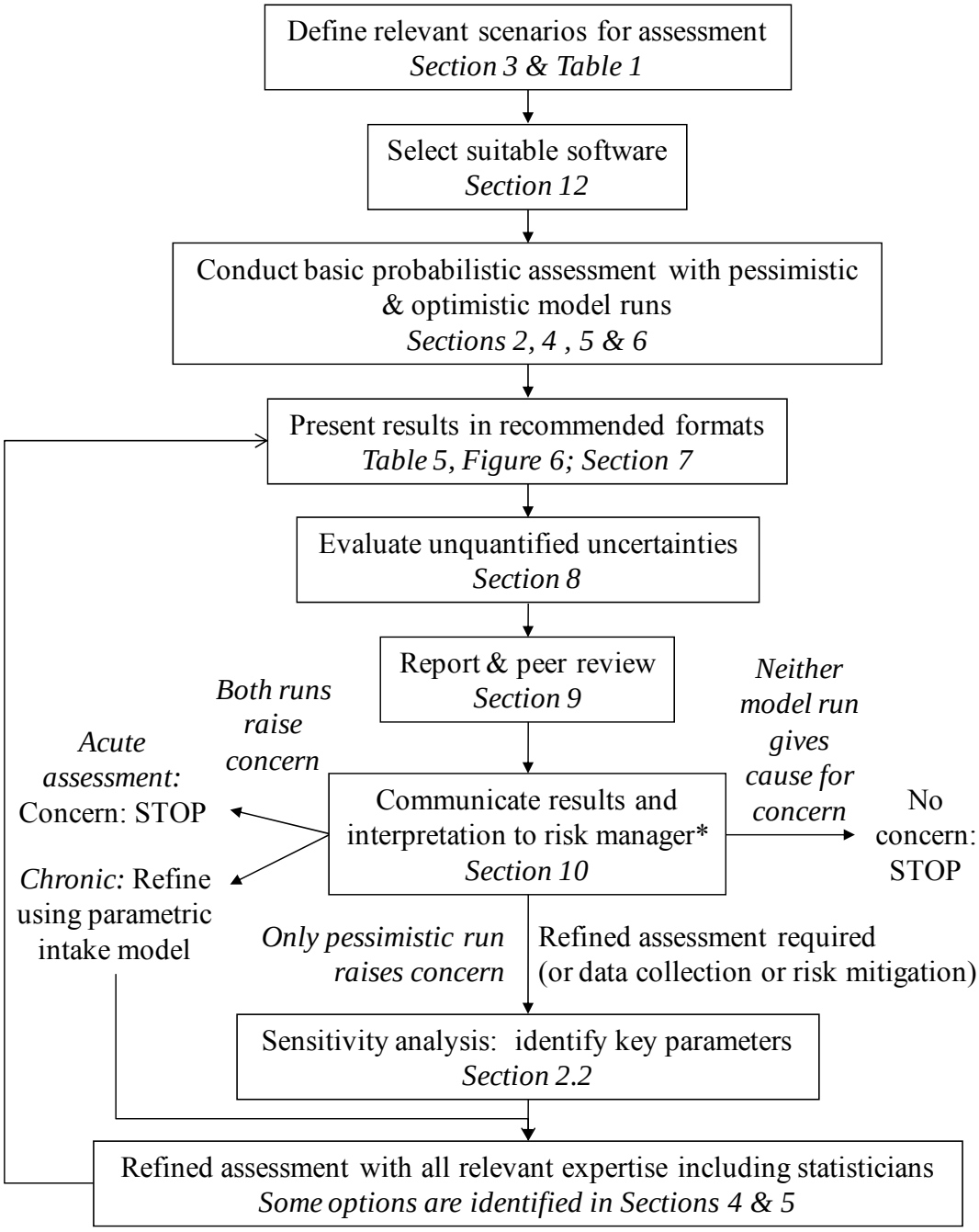


Figure 1: Summary of the main steps of the recommended approaches for probabilistic exposure assessment, with references to the relevant sections of this document. *In practice risk assessors may choose to refine the assessment to some degree before consulting with risk managers.

3. Problem definition

An essential first step in any exposure or risk assessment is to clearly define the purpose of the assessment and the specific question(s) and scenario(s) to be addressed.

The Terms of Reference identify three different contexts for dietary exposure assessment within the regulatory process for pesticides: authorisation of plant protection products; MRL-setting; and assessment of actual exposure based on residue monitoring data.

In practice, the last of these occurs in two different situations: first, in relation to individual cases of residues exceeding the MRL (referred to in this document as 'high residue events') to inform decisions on the need for enforcement actions; and second, in EFSA's annual reviews of monitoring data as required by Article 32 of Regulation 396/2005. There is also a distinction in Regulation 1107/2009 between authorisation of plant protection products, which occurs at National level, and the approval of substances, which occurs at EU level.

Figure 2 illustrates how these different situations fit into the overall sequence of events for authorisation, use, monitoring, review and enforcement. This is helpful in identifying the different types of exposure assessment that may be required (see below).

Figure 2 includes arrows in both directions between MRL-setting at EU level and Authorisation of Products at National level. This is because where a new use considered at National level requires modification of an MRL, this has to be assessed at EU level.

As indicated in Figure 2, a small proportion of residues exceeding the MRL occur in the marketplace (e.g., EFSA 2010a), even though the MRL is a legal limit. These residues are critical for risk assessment, so it is important to understand how and why they occur.

A proportion of the residues that are generated by use of an authorised product may be expected to exceed the MRL, even when there is compliance with the conditions of use. This is because the methods used for calculating MRLs are not aimed at identifying an absolute upper limit: rather they aim to produce a conservative estimate of the 95th percentile of the underlying residue distribution[17]. However, at least some of the residues that exceed the MRL are caused by unauthorised uses (e.g., EFSA 2010a).

One purpose of monitoring programmes is to identify where lots or consignments of a commodity in the marketplace contain residues above the MRL so that enforcement action can be taken to remove them. However, only a small proportion of all lots of a commodity is monitored, so the majority of those batches that have mean residues above the MRL may remain unidentified on the market.

In principle, a confirmed finding of residues above the MRL must lead to a removal of the sample from the market. The detailed procedure for confirmation varies between Member States but, in general, enforcement is considered when the residue found by an official laboratory exceeds the MRL by some specified margin, normally double to allow for measurement error of ±50% (SANCO, 2011a). This is intended to provide confidence (at the 95% level) that the measured value is actually above the MRL, and has not been over-estimated due to measurement uncertainty. In some Member States, additional samples are tested for confirmation. When considering the need for enforcement action, exposure assessment is sometimes used to check what level of risk is posed by the observed level of residues (as indicated in Figure 2).

[17] The OECD MRL calculator was adopted in February 2012 as the default method for calculating MRLs in the EU. The white paper on it states that its statistical goal 'in common with previous methodologies', is to produce a MRL proposal in the region of the 95th percentile of the underlying residue distribution, which is conservative in the sense that it will have a much greater propensity to make errors by overestimating the 95th percentile than by underestimating it for most datasets. (OECD, 2011, page 13).

In conclusion, there are a number of reasons why residues above the MRL are expected in the marketplace, and do indeed occur. This has been taken into consideration by the Panel in developing its recommendations for probabilistic modelling of exposure. The design of basic probabilistic exposure assessments depends on the context of the assessment, as summarised in Table 1 and explained in the following subsections.

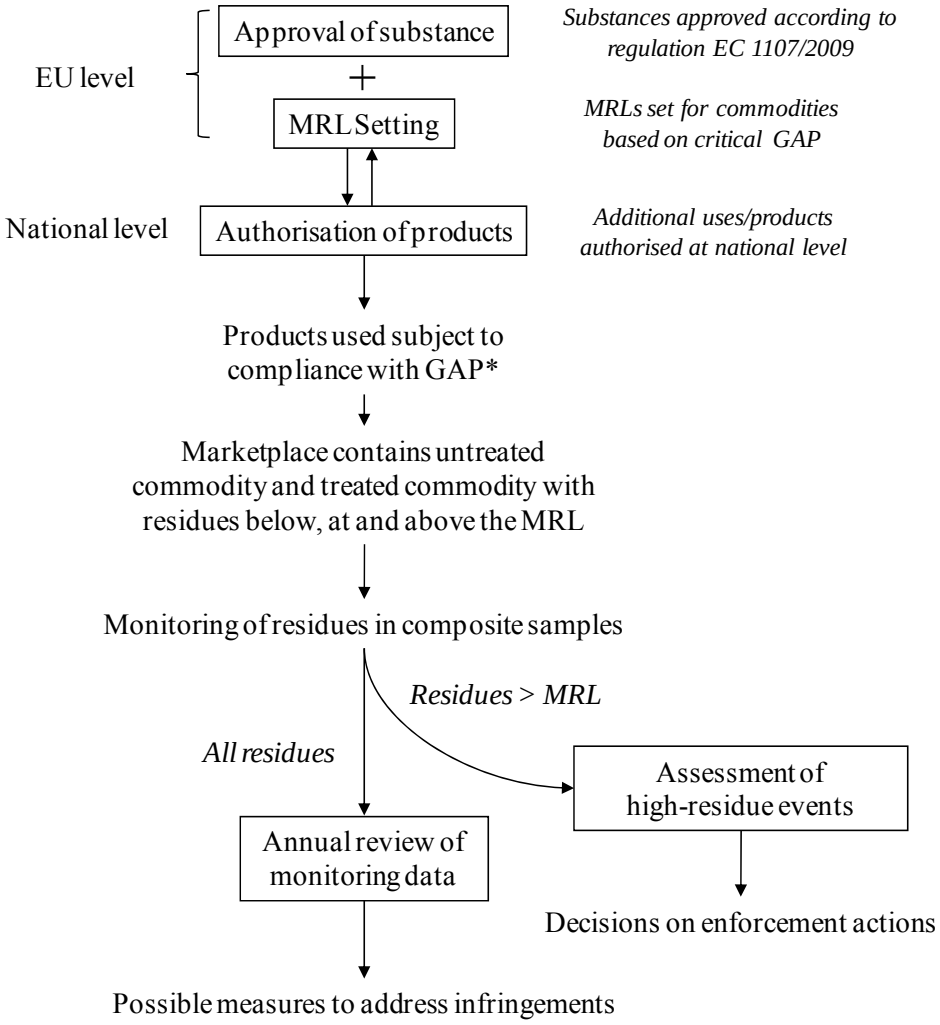


Figure 2: Illustration of regulatory contexts in which exposure assessment is required (indicated by boxed text). * GAP = Good Agricultural Practice.

Table 1: Impact of assessment context on major design elements of basic probabilistic dietary exposure assessment. Relevant metabolites and degradates should be included in all assessments. Unit-to-unit variability of residues should be included in acute but not chronic assessments. See text for more details.

Assessment context	Acute/chronic	Population	Focal[18] commodities	Primary data[19] for modelling residues	
				Focal commodities	Background commodities
Approval of substances	Acute & chronic	Whole of relevant population - OR - only persons/ person-days with consumption of focal commodities[20]	Commodity(ies) relating to the proposed uses	Supervised trial at critical GAP or feeding study with animal exposure at the critical GAP	Monitoring
MRL-setting					
Authorisation of products					
Annual review of monitoring data	Acute & chronic	Whole of relevant population	All	Distribution based on monitoring data	
High residue events	Acute	Only person-days in which the food in question is consumed	Single lot of commodity in which high residue occurred	Observed high residue	Monitoring
High residue events	Chronic	No assessment needed (chronic exposure not relevant for a single lot of commodity)			

3.1. Acute and chronic exposures

Consistent with general practice in dietary exposure assessment, acute exposures are calculated over a period of one day[21]. In principle, chronic exposure should be assessed as the average daily exposure of an individual over their lifetime. In practice, averaging may be applied over the duration of the survey providing the consumption data (in empirical modelling) or over an indefinite period (in parametric modelling, see Section 5). In addition, other patterns of exposure should be considered if there is a possibility of periods of exposure above the long-term average that might have toxicological significance (Renwick et al. 2003); however, this would require non-standard modelling approaches.

Both acute and chronic assessments can be relevant for all assessment contexts except high residue events, where only acute assessment is relevant (Table 1). This is because each high residue event relates to a particular lot of commodity in which the high residue has been found, and it is unrealistic

[18] The focal commodity is the commodity to which the approval, MRL, authorisation or high residue event relates. All other commodities in which residues of the substance may be present are referred to as background commodities.

[19] Table 1 shows primary data: where this is not available, or in refined assessments, other options may apply (see sections 4 and 5).

[20] For explanation of these options, see section on Population and individuals to be included.

[21] Shorter periods than 1 day may be justified for some types of chemical, but this involves special considerations and should be considered as a refined assessment.

to suppose that the same person will eat for chronic periods food from the same lot, or from different lots all with measured residues exceeding the MRL.

3.2. Population and individuals to be included

Dietary exposure assessments could in principle consider the whole of the EU or national population relating to the marketplace for which a use or MRL is authorised. However, both Regulation 396/2005 and 1107/2009 require that particular attention be paid to protection of vulnerable groups including pregnant women/unborn children, infants and children. This could be addressed either by conducting specific exposure assessments focussed on one or more vulnerable groups, or by assessing the overall population and displaying results separately for vulnerable groups. The choice of population(s) to be considered should therefore be defined in consultation with risk managers and may include the whole population, or specific subpopulations of interest.

In assessments for annual reviews of monitoring data, all individuals in the relevant population or subpopulation should be included, including non-consumers.

Acute assessments for high exposure events should consider those individuals who will consume the food in question[22], that is, the specific lot in which the high residue was found. To achieve this, the assessment should include only those person-days on which the focal commodity is consumed.

For approval of substances, MRL-setting and authorisation of products, assessment could in principle address either the whole of the relevant population or subpopulation, or only that part of the population who consume the commodities in question. Therefore, both options are included in Table 1 and either could be chosen, in consultation with risk managers. However, as the legislation requires that residues consequent on pesticide application 'shall not have any harmful effects on human health' (Regulation 1107/2009, Article 4.2), i.e., no harmful effects at all, it would be better from a technical perspective to restrict assessment to individuals or person-days where the focal commodity(ies) is/are consumed because this will provide a better picture of the upper end of the distribution of potential exposures (for any given size of simulation) than if the whole population is included. When reporting and interpreting the results, it is essential to make clear which population and individuals have been included.

3.3. Types of commodities and foods to be considered

Assessment of dietary exposure should include consideration of all plant and animal commodities in the form they are consumed (raw and/or processed) when they are expected to contain residues of the pesticide in question, and all foods that contain those commodities.

In assessments for approval of substances, MRL-setting and authorisation, a distinction is made between the **focal commodity**, to which the approval, MRL or authorisation relates, and all other commodities in which residues of the substance may be present, which are referred to as **background commodities**. It is necessary to include background commodities as these contribute to the total dietary exposure which is what determines the 'risks of the ADI or ARfD being exceeded' and whether 'any harmful effects' will occur[23,24]. The data available for modelling residues generally differs between focal and background commodities (see next section). If new uses for more than one

[22] Assessing exposure for consumers of the food in question is appropriate to inform decisions about 'suspension of the placing on the market or use of the food in question', one of the measures specified in Article 53 of Regulation 178/2002 (referred to by Article 35 of Regulation 396/2005).

[23] The level of exposure that might cause harmful effects will generally be expected to be higher than the ADI or ARfD, as these incorporate uncertainty factors that are intended to be protective.

[24] Residues in background commodities are not included in deterministic assessment, presumably because it is considered that the assumptions made about the focal commodity are sufficiently conservative that the background can be ignored. However, more realistic assumptions are used in probabilistic assessment, so the contribution of the background needs to be included.

commodity are being considered at the same time, then a single assessment should be done in which all the commodities affected by the new uses are treated as focal commodities, with other commodities as background.

In assessments of high residue events, the commodity in which the high residue has been found is the focal commodity and all other commodities in which residues of the substance may be present are considered as background commodities.

In assessments for annual review of monitoring data, all commodities are considered in the same way and no distinction is made between focal and background commodities.

Rarely eaten foods may not be represented in the consumption surveys that are available for assessment. Adjusting the modelling to take account of this is a complex challenge that is not appropriate for a basic probabilistic assessment. Any contribution of such foods to exposure will consequently be omitted from estimates produced by basic probabilistic assessments, unless it can be considered as a dietary alternative for other foods that are recorded[25]. Therefore, assessors should always check if the consumption survey includes data for all commodities with approvals or positive residues. If there are unrecorded foods, then the assessor should consider, as part of the evaluation of uncertainties, whether the contribution of those foods (based on a simple deterministic estimate) could be sufficient to materially alter the outcome of the assessment. If so, modelling adjustments to include those foods could be considered as an option for refined assessment.

3.4. Pesticide residues

The different types of assessment context (Table 1) require different treatment of residues, as described in the following paragraphs.

In assessments for approval of substances, MRL-setting and authorisation, residues in the focal commodity or commodities must be modelled using data from supervised trials/feeding studies, as the new uses under assessment will not yet be reflected in monitoring data. The supervised trial/feeding study data should be used to model the whole distribution of residues expected to result from the use. The whole distribution must be taken into account to meet the requirements of the respective legislation. Article 4.2 of Regulation 1107/2009 states 'the residues of the plant protection products, consequent on application consistent with good plant protection practice and having regard to realistic conditions of use…shall not have any harmful effects on human health'. Since the residues consequent on application vary, it is necessary to take account of the whole distribution in order to assess whether any harmful effects might occur. Article 29 of Regulation 396/2005 requires 'an assessment of the risks of the ADI or ARfD being exceeded as a result of the modification of the MRL'. For a probabilistic assessment of acute risks it is necessary to consider the whole distribution of residues that will occur after modification of the MRL, including values both above and below the MRL itself[26]. The current version of the Uniform Principles[27] requires estimation of 'the potential exposure of consumers' (Part B: Evaluation, paragraph 2.4.2.5), which again implies consideration of the full distribution of residues.

[25] For example, a consumption survey may contain no records of people eating goji berries, but if they are consumed as a dietary alternative (i.e., eaten instead rather than in addition) to other berries that are reported in the survey (e.g., blackberries) then modelling exposure via the other berries may be sufficient.

[26] A small proportion of residues exceeding the MRL is expected because the methods for calculating MRLs aim at a conservative estimate of the 95th percentile of the underlying distribution (OECD, 2011). This expectation is confirmed by monitoring data (e.g., EFSA 2010). Note that the current debate about whether to use the MRL or HR (highest residue in a supervised trial) in the IESTI equation relates only to deterministic assessment.

[27] Council Directive 97/57/EC of 22 September 1997 establishing Annex VI to Directive 91/414/EEC concerning the placing of plant protection products on the market. Official Journal L 265, 27/09/1997 pp. 0087 – 0109.

As well as considering the residues foreseen in the focal commodities, it is necessary to include also the residues present in background commodities as these contribute to the total dietary exposure (see preceding section). The modelling of residues in background commodities should be based on monitoring data as far as possible, but for commodities with authorised uses or import tolerances[28] that may contain residues but have too few measurements (or none), it will be necessary to make estimates based on extrapolation from monitoring data for other relevant commodities or based on supervised trial/feeding study data (see section 4.2.6 and, for consideration of untreated commodity, section 4.2.7).

Assessments for the annual review of monitoring data should include 'an analysis of chronic and acute risks to the health of consumers from pesticide residues' (Article 32, Regulation 396/2005). These risks depend on the full distribution of exposures occurring in the relevant population from all the commodities that may contain the substance, not just the particular residues found in the small proportions of those commodities that are monitored. As far as possible, these distributions should be based on the monitoring data supplemented by extrapolation or supervised trial/feeding study data for commodities with insufficient monitoring data.

In assessments of high residue events where residues exceeding the MRL have been found in the marketplace, the population considered is the group of people who consume the food in question, that is, the specific lot in which the high residue was found. To achieve this, the assessment should model the distribution of residues expected in that lot. This will vary above and below the measured value due to the combination of sampling variation (the measured value may be above or below the true mean of the lot) and unit-to-unit variability (individual units of a commodity will vary above and below the true mean). Again, it is necessary to take account of residues in background commodities consumed by the same people, as these contribute to determining the risk.

Note that assuming all of a commodity contains residues at the levels found in supervised trials/feeding studies will generally lead to over-estimation of exposure. On the other hand, using monitoring data implies an assumption that current or future levels of use are similar to those during the period to which the monitoring data relate, which might cause either over- or underestimation of exposure. These complications are taken into account in the approaches recommended in Sections 4 and 5.

Measured residues generally relate to a raw commodity. Food as eaten by the consumer comprises partly of raw commodity (e.g., raw apples), and partly of prepared foods (e.g., apple pie). The residue levels in prepared foods are influenced by several factors, including the composition or recipe for the food (e.g., apple pie is partly apple, and partly pastry) and processing effects (e.g., cooking). The effects of these factors may be taken into account using appropriate methods (see Sections 4 and 5). A special complication arises in assessments for high residue events, because consumers will sometimes (perhaps rarely) consume commodity from the lot in question in both raw and processed form (e.g., apples from the same purchase might be consumed raw and also after juicing, pureeing or cooking). This is taken into account in Section 4.

3.5. Residues of the same substance from veterinary drugs or biocides

Article 14 of Regulation 396/2005 requires that decisions on setting, modification or deletion of an MRL should take account of the possible presence of pesticide residues arising from sources other than current plant protection uses of active substances, when the methods to assess such effects are available.

Residues arising from veterinary and biocidal use of a substance which also has approval as a plant protection product will be included in monitoring data for food, and will therefore be addressed in the

[28] An import tolerance is an MRL set for imported products (see Glossary for definition).

assessments for annual review of monitoring data and also in the background commodities in other assessments (Table 1). Where monitoring data are not available at national or EU level, the higher of the veterinary MRL and the plant protection MRL (adjusted to the residue definition for risk assessment by the relevant conversion factor) could be used as a worst case estimate for residues.

Transfer of residues on preparation surfaces[29] to food can occur but is not normally considered in EU pesticide assessments[30]. If it was considered that this might contribute significantly to risk, it should be included in the evaluation of unquantified uncertainties in basic assessments and, if appropriate, accounted for quantitatively in refined assessment.

3.6. Metabolites, degradates and other transformation products

Metabolites, degradates, or other transformation products (hereafter collectively referred to as "metabolite/degradate") that significantly contribute to the dietary risk (complying with the residue definition for dietary risk assessment) should be included in the dietary exposure assessment (OECD, 2009). For each metabolite/degradate that is considered to contribute significantly to the risk, two factors must be addressed: 1) the potential for exposure to the metabolite/degradate in the human diet; and 2) the toxicity of the metabolite/degradate relative to the parent compound.

Only those metabolites and degradates identified as relevant in the regulatory assessment need be considered in probabilistic assessment. Where residues are quantified according to the residue definition for risk assessment, this will take account of relevant metabolites and degradates. In other cases, residues quantified according to the residue definition for monitoring should be adjusted to the residue definition for risk assessment. For basic assessments, appropriate factors for conversion of data to the residue definition for risk assessment are listed in the EFSA conclusion reports on peer-reviewed substances, and these should also be applied to the MRL when this is used in modelling. For refined assessments, more sophisticated methods could be an option (see Section 4.2.1). Further guidance on the treatment of metabolites and degradates is discussed by the Panel in a separate opinion (EFSA Scientific Panel on Plant Protection Products and their Residues (PPR), 2012).

3.7. Cumulative exposure to multiple substances

Both Regulation 1107/2009 (Article 4) and Regulation 396/2005 (Article 14) require that account should be taken of known cumulative and synergistic effects where scientific methods to assess such effects are available. All of the methodology described in this document for assessing dietary exposure to single substances is also relevant for assessing cumulative exposure to multiple substances via food. Guidance on additional methodology needed for assessing cumulative exposure to multiple substances is provided in Section 6. Guidance on methodology for identifying which substances should be considered together in assessments of cumulative exposure is being developed under a separate mandate to the Panel.

3.8. Exposure by routes other than food

The PPR Panel recognises that sources and routes of exposure other than food also contribute to overall risk, including drinking water, surface-to-hand transfers, and professional or residential exposure to pesticides.

The contribution of exposure via drinking water should be considered in probabilistic dietary exposure assessment. Methods for this are described in sections 4 and 5.

[29] This could include transfer between commodities prepared on the same surface, or transfer of substances used for treating or cleaning the preparation surface.

[30] The Dietary Risk Assessment Working Group (DRAWG) of the Technical Meeting on Biocides will in the near future publish guidance on how transfer of residues from preparation surfaces and food equipment should be dealt with in biocides assessments.

Article 33 of Regulation 396/2005 states that support measures relating to harmonised pesticide MRLs to be established at Community level shall include developing and using methods of assessing aggregate effects, which may imply taking account of the aggregation of dietary exposure with exposure of the same individuals via other routes (e.g., inhalation and dermal exposure), during activities as operators, workers, residents or bystanders. However, this requires further research and development before being addressed in a guidance[31].

4. Modelling acute exposure

Acute dietary exposures should be estimated for time periods of one day, for the scenarios indicated in Table 1. Acute exposure for the same individual varies between days due to day-to-day variation in consumption by individuals and unit-to-unit variation of residues in the foods they consume.

The basic inputs required for modelling dietary exposure are the amounts of pesticide *residue* that is present in and on foods and the *types and amounts of those food consumed* in a person's diet. However, a number of additional variables are also used. Some of these are adjustments required to allow the assessment to be conducted with the types of data that are normally available, while others allow the user to take account of factors that may modify exposure. They include:

- food conversion factors, to convert composite food products as recorded in dietary surveys (i.e., as eaten) to their individual raw agricultural commodities (RACs) or these commodities in the forms for which monitoring data are available;
- unit weights, required in acute assessments to divide weights of foods recorded in dietary surveys into individual items for some commodities (e.g., apples), so that between-unit residue variation can be modelled;
- variability factors, or other measures of the variation of residues between individual items of commodities;
- processing factors, to take account of changes in nature and amount of residues during the processing of raw agricultural commodities or commodities as monitored into processed commodities or ingredients (including peeling, juicing etc.); and
- estimates of the percentage of each commodity that is treated with the pesticide under assessment, for use in conjunction with supervised trial/feeding study data in optimistic basic assessments and in refined assessments.

Note that it is important to ensure that, for each ingredient of each food as eaten, the food conversion and processing factors are compatible with each other and do not double-count either the conversion or processing effects (see section 4.3).

The following sections discuss the possibilities for how each element of the acute exposure model could be handled in a probabilistic assessment and the difficulties that arise (e.g., due to limited data). They also explain the Panel's conclusions on which options should be used in optimistic and pessimistic runs for a basic probabilistic assessment, and which of them might be options for refined assessment. These conclusions are summarised in Table 2.

[31] For example, the EU Framework 7 research project Acropolis (www.acropolis.eu) is developing approaches for aggregate exposure assessment.

Table 2: Summary of recommended approaches for acute dietary exposure assessment (see the indicated text sections for detail). Each aspect applies to all types of assessment scenario (see Table 1) except where otherwise stated. The Panel expects that many of the approaches for basic assessment are likely to become available to users as built-in options in probabilistic software. Key to probabilistic methods in basic assessment (not shown for refined assessment): * fixed value; § variability modelled empirically, §§ variability modelled parametrically, ¶ uncertainty modelled by bootstrapping, ¶¶ uncertainty modelled parametrically.

Assessment component	Basic assessment		Options for refined assessment include:	Section no.
	Optimistic	Pessimistic		
Modelling food consumption	Empirical + bootstrap §¶	Empirical + bootstrap; examine which commodities contribute to upper tail exposures §¶	Parametric modelling	4.1.2
Water consumption	Zero	Treat as food if included in dietary survey§¶, or use deterministic estimate from drinking water assessment.*	More sophisticated estimates (see 4.1.4)	4.1.4
Separation of within & between individual variation of exposures	Not done		Parametric modelling (if separation is required)	4.1.2
Food conversion factors (recipes)	Use available recipe databases*		Quantify variability and uncertainty for foods driving exposure	4.1.1, 4.1.3
Unit weights	Use same values as in deterministic assessments*		Quantify variability and uncertainty for foods driving exposure	4.1.3
Residue definitions	Use residue definition for risk assessment, applying conversion factor where appropriate*. Evaluation of unquantified uncertainties.		Consider more sophisticated methods (see 4.2.1)	4.2.1
Residue measurement uncertainty	Not modelled.	Not modelled.	Consider including if thought potentially important.	4.2.11
Between lot/sample variation of residues	Empirical §	Lognormal for positive values (if n>2) §§	Parametric mixture models; extreme value models	4.2.3
Sampling uncertainty for lot/sample residues	Empirical Bootstrap ¶	Parametric for binomial & lognormal (if >2 positive values) ¶¶		4.2.5
Treatment of residues below LOR[32]	Treat as true zeroes*	Set <LOR to LOR*		4.2.7
Sampling uncertainty of proportion of residues below LOR	Empirical bootstrap ¶	Parametric model ¶¶		

[32] LOR: Limit of Reporting.

Percent crop treated (when using supervised trials/feeding studies data)	Approximate estimate of % crop treated*	Assume 100% of crop treated*	Refined estimate of % crop treated and the uncertainty of this	4.2.7
Limited amounts of monitoring data	Use only the data available for the commodities and country in question	Use appropriate data from other countries, other commodities or supervised trials/feeding studies	Future options might include extrapolation between substances	4.2.6
Residues from animal feeding studies	Estimate 4 or 8 values from 1x dose result, or set to zero*	Estimate 4 or 8 values from 1x dose result, or set to MRL or default MRL (0.01)*	More sophisticated estimates (see 4.2.1)	4.2.1
No supervised trials/feeding studies (as substitute for monitoring data)	If no trials or monitoring data, assume no residues.	Use appropriate trials data from other commodities or MRL level	Future options might include extrapolation between substances	4.2.6
Residues for non-authorised use	Treat as for authorised uses	Treat as for authorised uses except set <LOR to zero*	Treat as for authorised uses	4.2.1 and 4.2.7
Mean residue of focal commodity in high residue event	Set equal to mean of measured value(s) for high residue event*	Model uncertainty due to sample size and apply unit variability model §§¶¶	Model uncertainty due to sample size and apply refined model for unit variability	4.2.4
Between unit variation (e.g., variability factors, VF, or coefficient of variation CV)	None – unit residues all equal to lot/sample mean	Beta or Lognormal –conservative VF or CV (*simulation studies are needed to finalise the approach for this, see 4.2.9*) §§	Refined model, CV or VF varies between lots/samples, include correlation with lot/sample mean	4.2.9
Residues in prepared foods	Assume purchased, no unit variability	Assume prepared from the same sample of raw commodity as any raw consumption, and include unit variability §§	Refined assessment based on data or expert judgment	4.1.1, 4.4
Processing factors	Value used in deterministic assessment*	Set to 1 (no change) or use highest individual measured value, whichever is highest*	Quantify variability and uncertainty using data and/or expert judgment	4.3
Cumulative assessment	See section 6			6
Residues in water	Zero	Assume legal limit (0.1ppb for single substance, 0.5ppb for cumulative assessment)*	Consider using monitoring data from water	4.5
Unquantified uncertainties	Optional	Evaluate using uncertainty table	More sophisticated evaluation or quantification	8

4.1. Consumption

4.1.1. Data organisation and adjustment

As the outcome of probabilistic exposure assessment is to be compared to toxicological reference values which are expressed on a body weight basis (e.g., mg pesticide/kg bodyweight), exposure must similarly be expressed in relation to body weight so that these two quantities can be properly compared and evaluated. Therefore, consumption data should be linked to body weights for the same individuals, where possible.

Dietary consumption surveys collect data on foods "as eaten" (e.g., pizza, hamburger, beef stew) and not on their component parts (i.e., ingredients) and pesticide residue monitoring programmes generally collect residue data on raw agricultural commodities (e.g., apples, oranges, milk, etc.). Therefore, it is necessary to translate consumption of prepared foods from an "as eaten" food basis to a food commodity basis. This conversion is generally achieved using standard recipes which can be a part of the probabilistic dietary exposure software. More information on this conversion process is available in the Panel's previous Opinion on cumulative risk assessment (EFSA, 2008).

It is necessary to identify those prepared foods that can potentially be prepared at home from raw commodities. This is necessary because consumers will sometimes consume part of a single purchase of a commodity raw, and part processed. If that purchase of the commodity happens to contain above average-residues, the consumer will experience higher exposure than if they had purchased the prepared food separately or prepared it from a separate purchase of the raw commodity. This will occur sometimes though not frequently, e.g., a person who purchases apples with above-average residues and consumes some raw and some after juicing. Although this may be infrequent, it might be an important cause of upper tail exposures. Therefore, in *pessimistic model runs* for **basic probabilistic assessments**, it will be assumed that all meals of such foods are prepared at home and that if the same individual consumes the raw commodity on the same day, both will come from the same sample or lot (see section 4.3.1 for more detail on how this can be implemented). In *optimistic model runs*, it will be assumed that all prepared foods are purchased. If these alternative assumptions have a substantial impact on the overall exposure estimates, then one option for refined probabilistic assessment might be to model the proportion of prepared foods that is prepared at home in a more refined way based on appropriate data or expert judgment, if available.

In order to take account of unit-to-unit variation in residues in acute assessments (see section 4.2.9), it is necessary to divide the daily consumption of food items by the same person into individual units (e.g., convert "300g of apples" into an appropriate number of individual apples) using unit weights. Normally, this is done only for commodities with unit weights exceeding 25g (JMPR, 2003).

Rarely eaten foods may be under- or over-represented in the consumption surveys that are available for assessment and it is not practical to adjust for them in a basic probabilistic assessment. Therefore, assessors should always check if the consumption survey includes data for all commodities with approvals or positive residues. If there are missing foods, then the assessor should consider, as part of the evaluation of uncertainties, whether the contribution of those foods (based on a simple deterministic estimate) could be sufficient to materially alter the outcome of the assessment. If so, modelling adjustments to include those foods could be considered as an option for refined assessment.

4.1.2. Modelling of consumption

For acute exposure dietary assessment, variation in consumption has often been modelled "empirically" using the actual observed consumption data as recorded in a dietary survey, rather than by fitting parametric models to the data. Generally, one estimate of acute exposure is produced for each person-day in the survey, and consequently the output of the assessment represents variation between person-days.

Because even a large survey includes only a sample of the individuals in the total population, consumption data is subject to sampling uncertainty and will not represent perfectly the true diet of the population. This may be addressed by bootstrapping, a random resampling technique for quantifying sampling uncertainty (Efron, 1993). The assessment is repeated multiple times, each time replacing the dietary records with a sample of the same size drawn at random, with replacement, from the observed dietary records. This indicates the degree of sampling uncertainty in the distribution, but will only generate values that occur in the observed data and omits other values (most importantly, higher values) that would be found if the whole population were surveyed. Therefore, it is essential to examine the consumption data underlying the upper tail of the exposure distribution, consider by expert judgement how much higher the true upper tail of consumption could credibly be (i.e., whether higher consumption is plausible for the foods that contribute most to exposure), and take account of this when evaluating unquantified uncertainties affecting the assessment (see Section 8).

A particular advantage of the empirical approach to modelling consumption is that it retains potentially complex patterns that are present in the dietary survey, especially correlations between consumption of different foods (e.g., cereal products and potato products are eaten together less often than would be expected from their individual frequencies in the diet, Breuninger et al. 2003). However, this requires that the sample survey be of sufficient size such that these correlations are adequately represented in the data. This is less likely to be true for combinations of less-frequently consumed foods (e.g., turkey and cranberries). To guard against under-estimation, it is important to identify the foods underlying the upper tail of the exposure distribution, consider by expert judgement whether unobserved but credible combinations of those foods might give rise to higher exposures, and take account of this when evaluating unquantified uncertainties affecting the assessment (see Section 8).

Parametric modelling is an alternative approach, which uses distributions fitted to the survey data and can estimate the frequency of extreme consumption events by extrapolating beyond the range of the observed data. Some recent parametric approaches also estimate correlations between foods and quantify uncertainty (e.g., Kennedy, 2010). However, parametric methods require assumptions about the shapes of the distributions and the form of correlations (e.g., linear/nonlinear), which are themselves very uncertain. Further research is needed on these approaches and they are not yet available in readily-accessible software. Therefore, the Panel recommends that empirical modelling of consumption is used in both *pessimistic and optimistic runs* for **basic probabilistic assessments** of acute exposure, subject to bootstrapping and examination of tail values as outlined above. Parametric modelling of consumption may however be considered as one of the options for **refined assessment**.

The empirical approach is also limited to estimating the proportion of person-days that exceed toxicological reference doses. If some exposures above the reference dose are expected, the empirical approach will not indicate how they are distributed between individuals, e.g., whether a few individuals experience repeated high exposures or whether these exposures are spread over a larger number of individuals. If this information is needed by risk managers, it would require a **refined probabilistic assessment** using a parametric approach and formal separation of within- and between-individual variation, similar to the parametric approaches used for modelling chronic exposure (see below).

4.1.3. Food conversion factors and unit weights

Food conversion factors are used to convert dietary survey records of foods as eaten into the corresponding weights of their constituent raw agricultural commodities: e.g., to calculate the weights of wheat, tomatoes and other ingredients used in producing a given weight of pizza. These factors are generally derived from manufacturers ingredient lists and/or recipe books. They are usually organised in large "recipe" databases, which group prepared composite foods into a limited number of types and do not distinguish variations within these (e.g., it may be assumed that all pizzas contain the same proportion of tomato). Food conversion factors are often an integral part of the model software (i.e.,

not open for modification by the user) and are rarely if ever treated probabilistically. Clearly, actual food conversion factors are both variable and uncertain, but to quantify this for all food types would be a major undertaking.

Dilution is an important factor for some foods e.g., beer and tea. If the residue data relate to the processed food then correction for dilution is not required. Otherwise appropriate conversion factors can often be found in recipe databases, e.g., for beer to hops or tea to tea leaves. If some conversion factors are lacking from the recipe database, then they could be added on the same basis as other factors in the recipe database were derived. The origin of the conversion factors used needs to be documented and justified.

Unit weights are used in acute dietary exposure assessment to divide portions of a commodity recorded in a single survey record into the appropriate number of individual units. This is necessary in acute assessments to allow the modelling of variation in residues between units (see later). Recommended fixed default values are unit weights used by EFSA for acute risk assessment of pesticide residues (EFSA PRIMo 2[33]). The source of these data has been described in EFSA's reasoned opinion on the potential acute and chronic risk to consumers' health arising from temporary MRLs (EFSA, 2007b).

The limited data available for estimating food conversion factors and unit weights make it difficult to quantify their variation and the associated uncertainty. When conversion factors are used for *optimistic* or *pessimistic model runs* in **basic probabilistic assessments**, the Panel recommends using the same estimates for these parameters as are used in deterministic assessments. Unit weights and food conversion factors may differ between countries and regions, and conversion factors may also vary between different food processes. The source of the values used should therefore be clearly documented and the limitations of these estimates and their potential impact on the exposure estimates should be considered as part of the evaluation of unquantified uncertainties (section 8). In cases where these uncertainties are considered large enough to potentially change the risk management decision, more sophisticated modelling and/or collection of data could be considered as options for **refined probabilistic assessment**, targeted on those foods that contribute most to exposure in the basic assessment.

4.1.4. Water consumption

Exposure via drinking water should be included in the *pessimistic model runs* for **basic probabilistic assessment**. Water consumption may be modelled in the same way as food consumption, if the dietary survey includes consumption of water and if water used in recipes and added to tea and other drinks made with tap water, the use of water for cooking, etc. can be quantified. If this is not possible, an appropriate default estimate for water consumption for the relevant subpopulation could be used instead (EFSA Scientific Panel on Plant Protection Products and their Residues (PPR), 2012). For the *optimistic model run* in a basic assessment, exposure via water could be assumed to be zero. More sophisticated treatment of water consumption could be considered as an option for **refined probabilistic assessment**.

4.2. Residues

4.2.1. Data organisation and adjustment

The main types of residue data used in dietary exposure assessment are obtained from monitoring programmes and supervised field trials/feeding studies. The exposure scenarios specified in problem definition determine which types of residue data are preferred for each commodity in each assessment although it will often be necessary to use supervised trial/feeding study data data as a substitute for

[33]http://www.efsa.europa.eu/en/mrls/mrlteam.htm

monitoring data when the latter are absent or limited (see rationale in Section 3 and specific guidance below).

Where monitoring data are used, they should be taken only from time periods and regions where the actual use pattern of the substance is considered representative of the time period and region to which the assessment refers. In general, uncertainty will be reduced by using all relevant data. Often there will be little or no monitoring data for at least some of the commodities relevant to an assessment: approaches for coping with this are discussed in section 4.2.6. If a large quantity of monitoring data is available for a particular commodity and the number of positive values is large, then consideration could be given to using only those for the time periods and regions closest to the focus on the assessment. However, if examination of these data suggests significant variation between years or regions, sufficient data should be included to be representative of the range of that variation. More data should also be added if sensitivity analysis shows that residues for the commodity in question are a major contributor to uncertainty in basic probabilistic assessment.

Some monitoring is targeted and therefore not representative of the overall distribution of residues in the marketplace: possibilities for addressing this are discussed below in section 4.2.9. Other complications are also possible, e.g., monitoring samples are generally taken from retail outlets and may not be representative for the same commodity used in manufactured foods (e.g., if there are differences in the pesticide regimes, or in the conditions and duration of storage before use). Any concerns about the representativeness of residue data should be considered as part of the evaluation of unquantified uncertainty, and investigated in refined assessments if it appears they could be important.

The majority of data on pesticide residues in food, whether from monitoring or supervised trials/feeding studies, are measured for composite samples[34] containing multiple units of the raw commodity in question (e.g., 12 apples).

Supervised trials/feeding studies should use the residue definition for risk assessment, but residues from monitoring are in most cases quantified according to the residue definition for monitoring. In the latter case, they need to be adjusted to the residue definition for risk assessment (OECD, 2009) to take account of toxicologically relevant metabolites and degradates. Conversion factors for converting monitoring data to residue definition for risk assessment are sometimes available (see section 3.5). In principle, one might expect the ratio of metabolite or degradate to parent substance to increase over time, as increasing amounts of parent are metabolised or degraded, unless the metabolites or degradates are themselves lost more rapidly. A large number of supervised trial data for captan examined by the Panel showed a negative correlation between the concentrations of parent and metabolite (unpublished). Such patterns could lead to underestimation of exposure, since factors for converting monitoring data to the residue definition for risk assessment are normally estimated from supervised trial data whereas monitoring data, to which those factors are applied, are collected at longer time intervals after pesticide application when the ratio of metabolite to parent may often be higher. These issues may be further resolved in specific guidance on the establishment of the residue definition for risk assessment to be developed as follow up to the opinion of the Panel on the evaluation of the toxicological relevance of pesticides metabolites for dietary risk assessment (EFSA Scientific Panel on Plant Protection Products and their Residues (PPR), 2012). Until more guidance is available, the Panel recommends that both *optimistic and pessimistic model runs* in **basic probabilistic assessment** should use residue definitions for risk assessment according to current practice for deterministic assessment, and consider the impact of this as part of the evaluation of unquantified uncertainties (Section 8). More sophisticated methods for modelling metabolite levels could be an option for **refined assessment**, when more guidance is available and where suitable data to support this are available.

[34] The term "composite sample" in this Opinion is equivalent to "laboratory sample" as used in EU Directive 96/23 on Official Control of Food Commodities, and is used here to refer to samples comprising multiple units of the commodity in question.

It is common that residue data contain a proportion of concentrations that are reported only as being below a given limit, which is referred to as the limit of reporting (LOR) in this document. The proportion of values below the LOR can be very high in monitoring data (e.g., >80%). In case of pesticide/commodity combinations for which there is no registered use in their region of production, monitoring results showing no detection should be treated as true zeroes. All other censored residue data should be addressed using the approaches described in Section 4.2.7 (below).

No residues are normally expected in commodities for which no use of the pesticide is authorised and no import tolerance exists. Where monitoring data however show unexpected residues in commodities for which use of the pesticide in question is not authorised, these data should be used, so that the assessment reflects the exposures experienced by consumers.

When monitoring data for a plant commodity are not available but that commodity may contain residues transferred from treatments of previous crops through the soil, the level of residues present in the commodity may need to be estimated using expert judgment.

Animal feeding studies are conducted at 3 dose levels, normally 1x, 3x and 10x the MRL for animal feed. In some cases, a fourth dose level below 1x may be included in the study to reflect a lower livestock dietary burden situation. The regulation requires one feeding study for ruminants and one for hens. This provides only one measurement at the 1x dose in each animal commodity, compared to 4 or 8 measurements from supervised trials with plant commodities. Therefore, in order to enable variability and uncertainty of residues in animal commodities to be modelled probabilistically in the same ways as in plant commodities, the following approach should be followed. First, calculate the ratio between the residue measured in the animal commodity at the 1x dose and the residue in the animal feed for the 1x dose. Second, calculate the corresponding ratio for the 3x dose. If the two animal:plant ratios are similar (implying that absorption and excretion mechanism are not saturated), multiply each of the 4 or 8 measured residues from the supervised trial for the foodstuff by the animal:plant ratio from the 1x dose, thus generating 4 or 8 estimated residue values for the animal commodity. Then use these estimated residue values in the *optimistic and pessimistic runs* of the **basic probabilistic exposure assessment** in the same ways as values from supervised trials for plant commodities are used (see following sections). If the animal:plant ratios at 1x and 3x are markedly different, or if data to perform the ratio calculations are unavailable, then use the MRL or default MRL for the animal commodity as a fixed value in the *pessimistic run* and set the residue equal to zero in the *optimistic assessment*. In **refined assessments**, assessors may propose and justify more realistic assumptions based on data or expert judgment on the percentage of the animal feed crop that is treated, animal diet, and information from metabolism and feeding studies.

4.2.2. Conceptual model for variation of residues

This section discusses the conceptual framework for modelling residue variation. Detailed guidance on model implementation is provided in following sections (4.2.3 to 4.2.9).

For acute exposure assessments, it is necessary to consider residue levels per unit of consumption (e.g., an apple) and variation of residues between units in the marketplace. However, the available data are typically composite means (e.g., a measurement on a well-mixed quantity prepared from 12 apples). Consequently, in measured data some of the unit to unit variation is averaged out. Therefore, to take proper account of unit-to-unit variation when modelling acute dietary exposure, two levels of residue variation have to be considered: variation between the observed data values and unit-to-unit variation.

Deterministic acute exposure assessments for commodities with a unit weight over 25g aim to deal with residue variation in a conservative manner by using a "high residue" (HR) derived from the composite sample values, multiplied by a fixed "variability factor" to represent the degree to which residues in individual units may exceed the mean residue of a composite (JMPR, 2003). The

variability factor (VF) is defined as the ratio between the 97.5th percentile and mean of the distribution of unit residues, and this procedure is meant to ensure that the composite sample residue used in a deterministic assessment is adjusted to account for the fact that the residue of interest in an acute assessment is a high end residue.

The PPR Panel has considered two alternative conceptual models for dealing with the two levels of variation in probabilistic acute exposure assessments:

- Lot-based model: this considers that a particular commodity in the market is divided into lots, from which composite samples are taken, and that residue variability can be divided into variability within and between lots.

- Sample-based model: this considers that a particular commodity in the market could be divided into samples of the standard size used in monitoring, taken by the same procedure as is used by sampling inspectors, and that residue variability can be divided into variability within and between samples.

Both models can be illustrated in the same diagram, although the details differ (Figure 3). Each lot or sample contains units with varying residues, represented by the distribution in the lower part of the figure. Both the mean and the variance of the unit residues differ between lots or samples, as illustrated by the two distributions at the top of the figure. This is consistent with the variation observed in measured composite residues (sample means), and the evidence from unit datasets that the variability factor is itself variable (EFSA, 2005a).

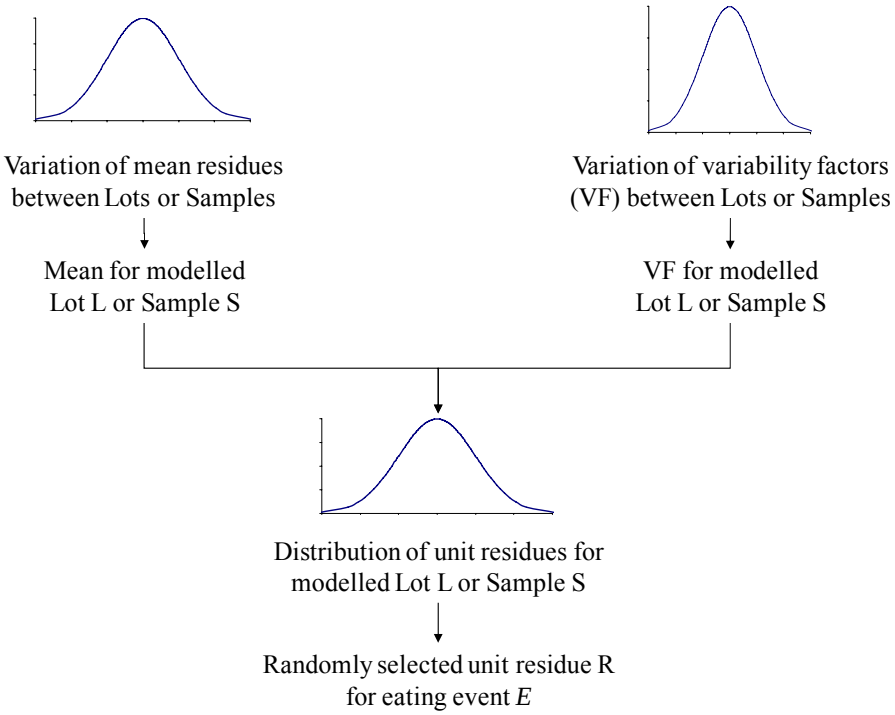


Figure 3: Graphical representation of conceptual model for variation of positive residues between lots or samples of the same commodity (upper left) and unit-to-unit variation of residues within lots or samples (lower graph). The proportion of residues below the LOR is modelled separately and their residues are set to zero or the LOR (see section 4.2.7). Note that in basic probabilistic assessments, the Panel recommends using alternative fixed values rather than a distribution (upper right) for the variability factor (see later).

An attraction of the lot-based model is that it attempts to reflect the real structure of the marketplace: commodities are traded in lots and consumers select units from lots. However, the definition of lots is not straightforward[35] and, in practice, some lots are mixed[36] and it cannot be assumed that each composite sample relates to a different lot. Another potential disadvantage is that monitoring data do not inform us about the exact lot means, but only provide estimated values with an estimation error that will depend on the variability of units within the lot. Due to these estimation errors the distribution of estimated lot means will be wider than the distribution of true lot means, potentially leading to an overestimation of exposure in the upper tail.

Attractions of the sample-based model include that it reflects directly the structure of the sampling process, which resembles the selection of units by consumers. The monitoring data inform us directly about the observed sample means, without sampling error. Therefore no confusion arises between between-sample modelling and within-sample (unit to unit) modelling. However, a degree of lot structure does exist in the real marketplace, and may influence the shapes of the distributions between and within samples.

[35] The Sampling Directive 2002/63/EC defines a lot as 'A quantity of a food material delivered at one time and known, or presumed, by the sampling officer to have uniform characteristics such as origin, producer, variety, packer, type of packing, markings, consignor, etc.' It also notes that: a) Where a consignment is comprised of lots which can be identified as originating from different growers, etc., each lot should be considered separately. b) A consignment may consist of one or more lots. c) Where the size or boundary of each lot in a large consignment is not readily established, each one of a series of wagons, lorries, ships bays, etc., may be considered to be a separate lot. d) A lot may be mixed by grading or manufacturing processes, for example.

[36] Data analysed by the Panel for a previous opinion (EFSA 2005a) show markedly multimodal distributions of unit residues in some lots.

A potentially important advantage of the sample-based model is that it implies a defined 'worst case' for the maximum unit residue, namely that all the residue is contained in one unit and is therefore equal to the measured residue for sample multiplied by number of units in sample. In this case, statistical distributions that are bounded between 0 and an upper limit, like the Beta distribution, can be used to model the unit to unit variation. In the lot-based model, there is no *a priori* value for the maximum possible unit residue, and statistical distributions with only a lower bound like the Log-normal distribution can be used to represent this.

In summary, coping with the complexity of the marketplace is challenging for both models, and neither is clearly preferable on theoretical grounds alone. Therefore, appropriate simulation studies should be conducted with different versions of each model to explore their performance in reproducing large datasets of measured unit residues. The Panel began work on some simulations but was unable to complete them before adoption of this Guidance. It is therefore recommended that appropriate simulations should be completed before making final conclusions on the approaches to be used for regulatory assessment (see Conclusions section).

Finally, it is important to note an important difference in the conceptual model for assessment of **high residue events**. This type of assessment is triggered by the reporting of a measured residue above the MRL in the marketplace (see section 3). In this case, a different conceptual model is required, focussed on consumers of 'the food in question', i.e., the particular lot of commodity in which the reported residue was measured. We refer to this commodity as the 'focal commodity' (see Table 1 in section 3). These consumers all take this focal commodity from the same lot, which has a single true mean residue[37], not a distribution of means as shown in the upper left part of Figure 3. Other foods eaten by these consumers come from multiple lots, so the conceptual model in Figure 3 does apply to the non-focal foods in high residue event assessments. The modelling of residues for this scenario is discussed further in section 4.2.4.

4.2.3. Modelling variation in residues between lots or samples

Like consumption, variation in residues may be modelled empirically, using the observed measurements, or parametrically by fitting a distribution to the observed measurements.

Empirical modelling will only generate residue values that appear in the measured data. This has the advantage that it does not generate higher values whose realism may be questionable. However, as residue datasets are normally small and the number of positive values still smaller, it will only generate a very small fraction of the values that actually occur, and their frequencies may differ widely from the true distribution due to sampling uncertainty. Bootstrapping provides an indication of the degree of sampling uncertainty, but the confidence intervals will only be reliable for large datasets and, even then, not in the tails. Most importantly, empirical modelling will almost always under-estimate upper-tail dietary exposures because values from the upper tail of the true distribution occur rarely and therefore are unlikely to occur in residue datasets of typical size. For this reason, for acute exposure, the Panel proposes to use empirical modelling only for the *optimistic model run* in **basic probabilistic assessments**. This applies equally whether the data are from monitoring or supervised trials/feeding studies. Where monitoring data are used, values below the LOR may be assumed to be true zeroes in the optimistic run. When trial data are used, in the optimistic run they may be combined with an estimate of the proportion of crop that is untreated (see Section 4.2.7).

Parametric modelling uses parametric distributions that are based on the observed data but generate additional values below, between, and above the observed values. This has the advantage of being able to represent the full range of potential residues, but requires assumptions to be made about the shape of the distribution. The limited size of residue datasets makes the choice of distribution shape very

[37] However, the true mean is uncertain, because it must be estimated from only one or a few residue measurements. See section 4.2.3 for further discussion.

uncertain, especially in the upper tail. If unbounded distributions are used (e.g., lognormal), they will certainly generate a small proportion of unrealistically high values, even if they fit the data well.

For the parametric approach, it will generally be necessary to use a combination of distributions to model residues for each commodity: a binomial distribution to represent the frequency of positive residues, combined with one or more distributions to represent the variation of positive residues within samples or lots. Residues reported as being below the LOR may be true zeroes (untreated commodity) or low positive residues. It is possible to model the proportion of true zeroes and the distribution of positive values together in a single statistical model that takes account of dependencies between them. However, these models require specialised methods and their performance for small datasets has not yet been examined so they are not suitable for basic assessment but may be considered as an option for **refined probabilistic assessment**.

For *pessimistic model runs* in **basic probabilistic assessment**, the Panel proposes that the proportion of residues below the LOR and the distribution of positive values should be modelled independently. The proportion of residues below the LOR should be modelled using a binomial distribution based on the observed proportion (see Section 4.2.7). There will often be too few positive values to discriminate well between alternative choices of parametric distribution so for basic assessment the Panel proposes using the lognormal distribution. This may often provide a reasonable fit within the range of the observed data, as Boon et al. (2003b) found for 10 pesticide-commodity combinations. Furthermore, a lognormal distribution is supported from a theoretical viewpoint in that residues are positive, positively skewed and originate from mechanisms generating the lognormal distribution under a variety of biological circumstances[38]. However, in using the lognormal distribution for the pessimistic run, it is not assumed that this distribution necessarily fits the data well. Instead, the intent is that assuming a lognormal distribution, quantifying sampling uncertainty for this and including unit variability will give an upper 95% confidence interval that has a high probability of being above the true distribution in the upper tail, which is the area of most interest for decision-making, without being excessively higher. The Panel began work on some simulations to examine whether this criterion is met, but was unable to complete them before adoption of this Guidance. It is recommended that appropriate simulations should be completed before making final conclusions on the approaches to be used for regulatory assessment (see Conclusions section).

If use of empirical modelling (in the optimistic run) or the lognormal distribution (in the pessimistic run) substantially changes which commodities contribute most to the estimated exposures or the magnitudes of their contributions, distribution fit could be examined as part of **refined probabilistic assessment**. The goodness of fit of the positive residues to the lognormal distribution could be evaluated for the most important commodities using visual examination of Q-Q or similar plots (e.g., Vose, 2008). Significance tests of distributional fit can be misleading, because their power depends on the size of the dataset. Where there are sufficient data to assess the goodness of fit, alternative distributions could be considered for use in the refined assessment, provided this is fully documented and justified and that appropriate methods can be implemented to quantify the associated sampling uncertainty. Where there are too few data to evaluate goodness of fit, it would be prudent to retain the lognormal distribution as a default and take the uncertainty about distribution choice into account when evaluating unquantified uncertainties (see section 8). In the future, consideration could also be given to more sophisticated options such as nonparametric modelling, extreme value theory and the pooling of data for multiple pesticides to model a shared distribution shape[39].

Many commonly used parametric distributions, including the lognormal, extend to infinity, and therefore their upper tails include values that are clearly unrealistic (e.g., concentrations over 1kg/kg).

[38] See the discussion of R-P (Random Product) processes the Theory of Successive Random Dilutions (SRD) in Wayne R. Ott's *Environmental Statistics and Data Analysis,* Lewis Publishers, 1995.

[39] These approaches show promise for modelling pesticide residues but require further evaluation and are not yet available in exposure modelling software (Paulo *et al.* 2006, Kennedy et al. in prep.).

Although extreme values will very rarely be sampled in probabilistic modelling, when this does occur it would be misleading for decision-making. Unfortunately, there is usually no good basis for choosing any specific residue value (other than the absolute maximum of 1kg/kg) as the upper bound for a truncated or bounded distribution. Therefore, the realism of residue values in the upper tail of the output from *pessimistic model runs* for **basic probabilistic assessments** should always be checked by examining 'drill down' statistics (see Section 7).

4.2.4. Modelling the mean residue in the focal commodity for a high residue event

This type of assessment is triggered by the reporting of a measured residue above the MRL and is focussed on consumers of 'the food in question', that is, the particular lot of commodity in which the reported residue was measured (see section 3). These consumers all take this focal commodity from the same lot, so the true mean residue in the focal commodity should be treated as a fixed value: although the true value is uncertain, it must be estimated from only one or a few residue measurements. How residues in individual units of the commodity vary around the mean value should be modelled using the approaches set out in section 4.2.9. The mean value itself may be set equal to the measured residue in *optimistic model runs* for **basic probabilistic assessments**. In *pessimistic model runs*, the sampling uncertainty for the true mean value should be modelled assuming the underlying distribution of mean residues between samples or lots is lognormal using the methods described in the following section. In **refined probabilistic assessments**, more sophisticated approaches to modelling the sampling uncertainty could be considered (e.g., using information from other commodities and substances to model the form and variance of the underlying distribution, see the following section).

4.2.5. Modelling uncertainty due to the limited size of residue datasets

The amount of monitoring data available varies widely between commodities and substances, but the number of positive residues per commodity is often very small, frequently as low as one or two values. Supervised trials normally provide eight measurements for plant commodities, or four in the case of minor crops. Animal feeding studies normally provide only a single measurement at the '1x' dose, which may be used to generate 4 or 8 estimated values using the method described in section 4.2.1.

Such small numbers of positive measurements cause high uncertainty in evaluating the shape of the full distribution and estimating its mean and variance. This is referred to as sampling uncertainty because it is caused by variation in the values obtained when samples are drawn from a population.

The influence of sampling uncertainty can be very large when the dataset is small, but is also present for large datasets, especially in the tails. Therefore, the potential magnitude of sampling uncertainty and its impact on the assessment outcome must be considered.

In *optimistic model runs* for **basic probabilistic assessments**, sampling uncertainty may be quantified by empirical bootstrapping. Bootstrapping is a computer-intensive methodology for quantifying sampling uncertainty (Efron, 1993). Briefly, the assessment is repeated multiple times, each time replacing the measured residues with a sample of the same size drawn at random, with replacement, from the measured residues. The multiple output distributions generated by the multiple runs are then used to estimate confidence intervals for the "true" distribution. The number of bootstrap iterations should be sufficient to generate stable confidence intervals (this should be checked by making three or more repeat calculations of the estimated confidence intervals and, if needed, increasing the number of bootstrap iterations). Confidence intervals obtained by bootstrapping will be very approximate when the resampled dataset is small, and also in the tails of the distribution even when the dataset is large. Sampling uncertainty is highest when there is only one observed value, but bootstrapping will not reflect this at all. Furthermore, empirical bootstrapping is limited to recombinations of the observed values, and cannot represent uncertainty about the existence of values outside the observed range. This is why it is recommended here only for optimistic model runs.

In *pessimistic model runs* for **basic probabilistic assessments**, sampling uncertainty should be estimated using parametric models, that is, models based on the sampling behaviour of an appropriate parametric distribution. This provides distributions for the uncertainty of the parameters of the distribution, based on the sample data. For reasons explained in section 4.2.3, the lognormal distribution is assumed as a default for the basic assessment. The logarithms of the residues are then assumed to follow a normal distribution for which an analytical solution for sampling uncertainty is available (e.g., Vose, 2008). The uncertainty of the standard deviation of the normal distribution is described by an inverse chi distribution, as follows:

$$\sigma \sim \sqrt{\frac{(n-1)s^2}{\chi^2_{(n-1)}}} \qquad \textit{Equation (1)}$$

Where '~' means 'distributed as', s is the standard deviation of the observed log residues, n is the sample size, and $\chi^2_{(n-1)}$ is the chi-square distribution with n-1 degrees of freedom.

The uncertainty distributions for the mean and variance are inter-dependent, so the distribution for the mean is considered after drawing a value for the variance using equation (1). This variance is then treated as known, so the uncertainty of the mean can be described by a normal distribution (rather than Student's t distribution), as follows:

$$Mean\log residue \sim Normal(\bar{x}, \frac{\sigma^2}{n}) \qquad \textit{Equation (2)}$$

Where $\bar{x}$ is the mean of the observed log residues and σ is the drawn value for the standard deviation.

The above method can be used when there are two or more different measured values. When there is only a single positive value, the observed value is taken as the estimate of the mean, but a sample variance cannot be calculated. Ignoring variation and uncertainty for such commodities would clearly be unconservative, which is acceptable for the optimistic model run but not for the pessimistic model run. Therefore, the most relevant available information should be used to estimate a surrogate standard deviation for commodities with only one positive value, e.g., monitoring data for another commodity where there is reason to expect a similar distribution of residues, or supervised trial data for the commodity in question or another for which extrapolation is appropriate. A surrogate standard deviation may also be used when there is more than one positive measurement for the commodity in question, but they all have the same value, so the sample standard deviation is zero (which can happen by chance, especially when measured values are rounded, but is not a realistic basis for modelling).

In **refined probabilistic assessments**, sampling uncertainty could be modelled parametrically, and consideration may be given to distributions other than the lognormal, where there is evidence to support them, and to using more sophisticated methods for improving the estimates by using additional information from other commodities and other substances.

4.2.6. Using residue data from different sources to increase sample size

Very small datasets have very high sampling uncertainty. Although this will be quantified using the methods described in the preceding section, it is desirable to reduce the uncertainty by using information from other sources, which is referred to as extrapolation. This is also desirable for commodities that lack any positive residue measurements, unless the number of measurements below the LOR is large.

The most obvious sources of information for extrapolation are:

- monitoring data for the same commodity in different countries where the residues are expected to be similar;

- monitoring data for other commodities that are expected to follow a distribution similar to that for the commodity in question;
- supervised trial data for the commodity in question; or
- supervised trial data for another commodity for which similar residues are expected.

Before using either monitoring or supervised trials data from other countries or commodities, their relevance for extrapolation to the commodity in question should be critically assessed. Extrapolation should only be considered for pairs of commodities listed in guidance document SANCO 7525/VI/95, and only when it can be reasonably expected that the use and usage practices of the pesticide in question are the same in both commodities. Similarly, extrapolation between EU countries should follow guidance provided by SANCO (2011b). Extrapolation from countries outside the EU should only be considered when it can be reasonably expected that residues will be similar in both countries. All extrapolation should be fully documented and justified in the assessment report.

In principle, consideration could also be given to using monitoring data or supervised trials for other substances in refined assessments, where there is justification to expect these to be similar.

Clearly, extrapolation of any type introduces additional uncertainty which must be taken into account. In a **refined probabilistic assessment**, extrapolation or combining of data from different sources should be done using appropriate statistical methods which quantify the associated uncertainty. However, such methods are not practical for basic assessments.

In *optimistic model runs* for **basic probabilistic assessments**, no extrapolation is necessary. The assessment may be conducted using only the residue data that are available. This will underestimate dietary exposure because it will assume residues are always zero in commodities that have no positive measured values and will ignore the possibility of residues higher than those observed for commodities that have positive values. However, this is acceptable in an optimistic model run.

In *pessimistic model runs* for **basic probabilistic assessments**, extrapolation should be used to enable parametric modelling of residues for every commodity that has an authorised use or MRL for the substance under assessment. The Panel recommends the following procedure.

- If there are 2 or more different monitoring values:
    - As a first step, assume a lognormal distribution and model uncertainty parametrically, as described above. Inspect the simulated residues underlying the upper tail of the exposure distribution using drill down outputs (see section 7).
    - If this reveals residues that the assessor believes are unlikely to occur even rarely, then look for data from other countries and/or commodities for which there is established extrapolation, and merge these with the monitoring data you have. Rerun the model as before, and inspect the simulated residues in the upper tail.
    - If using other country/extrapolation data still generates values the assessor believes to be unlikely to occur even rarely, rerun the model replacing the monitoring data with data from supervised trials/feeding studies (if there is more than one trial, use all that are relevant). If this reduces the upper confidence bound and generates more credible simulated residues, then use this in place of the upper confidence interval obtained with the monitoring data. Explain, when reporting the results, that this has been done, and discuss carefully its impact on the assessment.

  - If there are no supervised trials data for the commodity in question, substitute trials from other commodities for which extrapolation is accepted. Again, use all the supervised trials that are relevant.

  - If there are no usable supervised trials at all, then use the results obtained with only the monitoring data.

- If there are fewer than 2 different monitoring values, use supervised trial/feeding study data instead. For plant commodities, if there is more than one supervised trial, use all that are relevant. If there are no supervised trial data for the commodity in question, substitute supervised trials from other commodities for which extrapolation is accepted. Again, use all the supervised trials that are relevant. If there are no relevant supervised trial data at all, and less than two different monitoring values, use MRL itself as a fixed value but state clearly in the report that this has been done. For animal commodities, generate 4 or 8 estimated values from the feeding study as described in section 4.2.1 or, where this is not possible, use the MRL or default MRL in the pessimistic model run and zero in the optimistic run.

Data from supervised trials relate to treated commodities, whereas monitoring data generally include both treated and untreated commodities. Approaches for taking account of the percentage of crop that is treated are discussed in the following section.

4.2.7. Handling of untreated commodity and residues below the limit of reporting

Monitoring data based on composite samples frequently contain a high proportion of values below the Limit Of Reporting (LOR). Data from supervised trials may also contain values below the LOR.

In supervised trials, all the commodity is treated, so <LOR values are likely to represent positive residues below the LOR. Monitoring data relate to the marketplace, which generally includes both treated and untreated commodity, so some of the <LOR values may be low positive residues but others will be true zeroes (untreated commodity).

Various statistical methods are available for estimating values below a limit of reporting, and for modelling mixtures of positive values and true zeroes. Some of these methods were evaluated in a recent study by EFSA (2010b). It was concluded that the performance of the evaluated methods was questionable when the number or proportion of positive values was small, and on this basis it was recommended that probabilistic exposure assessment should not be conducted when there are less than 25 positive samples, or when more than 80% are censored (<LOR). In most pesticide assessments, these requirements will be met for only a few major commodities, such as apples. EFSA (2010b) suggest that, in such cases, similar food categories can be pooled together to obtain larger sample sizes, or additional data should be collected. Even when data are pooled, as described in the preceding section, there will still be many commodities that fail to meet the requirements proposed by EFSA (2010b). For many commodities, the proportion of crop treated may be less than 20%, so collecting further data will not meet the proposed requirement. However, probabilistic approaches are needed for cumulative assessments, and to take account of upper tail exposures in higher tier assessments for single substances. Therefore, in the case of pesticides, the PPR Panel proposes an alternative strategy using different assumptions in the optimistic and pessimistic model runs to take account of the uncertainty in a way that is practical for basic probabilistic assessments.

For *optimistic model runs* in **basic probabilistic assessments**, the following procedures should be used:

- When monitoring data are used, values below the LOR should be treated as true zeroes. Uncertainty about both the proportion of values below the LOR and the distribution of residues in treated commodity should be quantified by empirical bootstrapping of the full set of observations, both above and below the LOR.

- Supervised trial data are used in optimistic model runs only for the focal commodity in assessments for approval, MRL-setting or authorisation (see Table 1 in section 3). Values below the LOR may be treated as true zeros, as an optimistic assumption[40]. In addition, allowance may be made for the proportion of that commodity that is expected to be untreated, by adding the appropriate proportion of untreated values after bootstrapping the supervised trial data. In the optimistic basic assessment, the proportion of the commodity that is untreated can be an approximate judgment. Similarly, an appropriate proportion of the animal commodities can be considered as containing no residues on the basis of the same approach (untreated feedingstuffs). This may be represented by adding a number of zero values to the values generated from the feeding studies as described in section 4.2.1.

For *pessimistic model runs* in **basic probabilistic assessments**, the following procedures should be used:

- When monitoring data are used, values below the LOR should be replaced with the LOR as a conservative assumption. Uncertainty about the proportion of values below the LOR should be modelled parametrically (see below for method), and uncertainty about the distribution of residues above the LOR (ignoring those below it) should be modelled parametrically as described in section 4.2.5.

- If supervised trial/feeding study data are used for the focal commodity, values below the LOR should be replaced with the LOR, and it should be assumed that 100% of the commodity will be treated. The distribution of values should then be modelled parametrically as described in section 4.2.5 above.

- Where supervised trial/feeding study data are used as a substitute for few or no positive values in monitoring data (see section 4.2.6), uncertainty about the proportion of values less than the LOR for monitoring should be modelled parametrically based on the available monitoring data (see below for method), with monitoring values below the LOR for monitoring being replaced by that LOR. Values above the LOR for monitoring should be simulated from the supervised trial/feeding study data, by first replacing any trial values below the LOR for the trial with the trial LOR, and then modelling the distribution of positive values with uncertainty quantified parametrically (see section 4.2.5).

In both the optimistic and pessimistic model runs, if any values below the LOR relate to samples from regions of origin where there is no registered use for the pesticide and commodity in question, then those values may be considered as true zeroes. However, any positive concentrations recorded from regions without registered uses may result from illegal use and should be retained in the model.

In the steps described above, uncertainty about the proportion of values less than the LOR for monitoring should be modelled parametrically assuming a binomial distribution with true proportion p. The uncertainty of p may be modelled by the Beta distribution:

$$p \sim Beta(r+1, n-r+1) \qquad \textit{Equation (3)}$$

Where n is the sample size and r is the number of residues below the LOR[41].

If sensitivity analysis shows that the treatment of values below the LOR has a large influence on estimated exposures, consideration could be given to more sophisticated approaches in **refined**

[40] Values close to zero might be feasible for treated commodity in some conditions, e.g., where there is high plant metabolism and/or where there is a long period between treatment and harvest.

[41] This is the posterior distribution for a binomial proportion estimated from data, assuming an initial expectation (before seeing data) that the true proportion is equally likely to lie anywhere between 0 and 1 (i.e., a prior distribution that is uniform between 0 and 1; see page 234 of Vose, 2008).

probabilistic assessments. This might include exploring the capability of advanced modelling approaches (e.g., those discussed as possible areas for further work by EFSA, 2010b) to address the challenges posed by the limitations of the data available for the assessment in question. Refined assessment could also make use of estimates of the proportion of each commodity that is or will be treated, but this should be done more rigorously than in the optimistic basic assessment, based on the observed proportion of values below the LOR together with information on the existence of registered uses and the method, timing and extent of use in the region of origin[42]. Account might also be taken of information on prevalence of the target pest or disease of the pesticide and on factors affecting the market shares of alternative products for the same use. The use of all these types of information involves expert judgment, which should be fully documented and justified. Uncertainties affecting these judgments should be quantified using formal methods of expert elicitation (for an overview see O'Hagan et al. 2006) or considered as part of the evaluation of unquantified uncertainties (section 9).

Note that using monitoring data for modelling implies an assumption that current or future levels of pesticide use are similar to those during the period to which the monitoring data relate, which might cause either over- or underestimation of exposure. Uncertainty about this should be quantified using formal expert elicitation or considered as part of the evaluation of unquantified uncertainties (section 9).

4.2.8. Addressing non-random sampling in residue monitoring data

The random sampling methods normally used in probabilistic modelling assume that the residue data are representative of the variation of residues in commodities available to the consumer population under assessment. Deviations from this may occur through deficiencies in sampling design or due to targeted (or selective or probability-based) sampling of particular cropping practices, cropping seasons or region of production, or investigative sampling of suspect lots of commodity. Targeted sampling is expected in most cases to cause overestimation of exposure, but could also cause underestimation, as targeting of sources or seasons for one pesticide might cause under-representation for other pesticides, e.g., fungicides in winter vs. insecticides in winter.

EU coordinated monitoring is untargeted (see EFSA annual monitoring report, 2010a), but private monitoring and some national monitoring is targeted. In a database of pesticide residue concentrations, it is currently often not possible to distinguish data obtained from targeted sampling (although this is expected to improve in future), so the possible influence of targeting must be considered as part of the evaluation of unquantified uncertainties (section 8). If part of the data are clearly identified as targeted, they could be excluded from the assessment, but this must be justified. Alternatively, or when the nature of targeting is less clear, its impact could be tested by rerunning the basic assessment separately for subsets of the data that were collected in different ways. If the results differ significantly, one option would be to undertake a refined assessment using more sophisticated modelling methods to take account of the non-random sampling. For example, if detailed information on the nature of targeting and sampling design is available, this can be used to weight the data in an appropriate way.

4.2.9. Modelling unit to unit variability of residues

The residues data available for use in dietary exposure assessment generally relate to composite samples, not individual units of commodity. Therefore the measured values represent the average of a number of units and do not reflect the full range of variation occurring in individual units, which needs to be considered for acute assessments. In deterministic acute exposure assessments, unit to unit variability is represented by a variability factor (VF), defined as the ratio between the 97.5th percentile and mean of the distribution of unit residues (JMPR, 2003).

[42] Percentage of crop treated may be estimated well for countries that conduct detailed surveys of pesticide usage, but these are lacking in many countries. Furthermore, data on pesticide treatments for imported commodities are generally very limited.

For consistency, this guidance for probabilistic assessments also uses the variability factor for this purpose, although from a technical perspective it would be preferable to quantify unit-to-unit variability using the coefficient of variation (CV). This is because the interpretation of the VF depends on the distributional form assumed and, for some distributions including the commonly-used lognormal distribution, the same VF may refer to more than one level of unit-to-unit variability[43].

As explained in more detail in section 4.2.2, the PPR Panel considered two alternative conceptual models for dealing with these two levels of variation in probabilistic acute exposure assessments: a lot-based model and a sample-based model. Neither is clearly preferable on theoretical grounds alone, so the Panel considers that simulation studies should be conducted with different versions of each model to explore their performance in reproducing large datasets of measured unit residues. The Panel began work on some simulations, but was unable to complete them before adoption of this Guidance. The Panel therefore recommends that appropriate simulations should be completed before making final conclusions on the approaches to be used for regulatory assessment (see Conclusions section). For the meantime, both approaches are referred to in the following discussion of methodology for modelling unit-to-unit variability.

Both the lot-based and sample-based models are based on the conceptual model illustrated in Figure 3. When modelling the residue in a specific unit of commodity, a value is first drawn at random from the distribution of mean residues, which is based on composite residues from supervised trials or monitoring data (top left of Figure 3). The sampled value is then considered as the mean residue for the lot or sample from which the consumed unit is derived. To model the residue in the consumed unit, a value is drawn from the distribution of unit residues in this sample or lot (shown in the centre of Figure 3) which is specified by the sampled mean together with the variability factor for that sample or lot[44].

Following international discussions about the choice of default variability factors, the PPR Panel examined a large amount of residue data on single units from existing studies including both data from supervised field trials and from market surveys (EFSA, 2005a). The Panel found that the variability factor was itself variable, i.e., the degree of unit-to-unit variability differs between different studies, which seems reasonable given the existence of variation in pesticide properties, in crop characteristics, in application techniques and in the effects of harvesting, storage and transport. If this variation in the variability factor has a significant impact on exposures, then it should be considered in probabilistic modelling, as illustrated in the upper right side of Figure 3.

However, there are additional complications that need to be considered. First, EFSA (2005a) found that the distribution of variability factors differs between supervised field trials and market surveys: this is at least partly due to the fact that market samples may contain units derived from mixed lots which may include treated and untreated commodity and commodity with different treatment histories. This means that variability factors estimated from market samples are likely to overestimate the variability within a lot comprising exclusively of treated commodity, while variability factors from supervised trials probably underestimate the variability present in treated lots in the marketplace.

A second complication is that the variances of residues in different composite samples of the same commodity might be expected to correlate negatively with the mean residues of those samples. This is because a high mean residue is likely to occur in samples which, by chance, contain only units from the upper tail of the overall distribution, and therefore the variation between the units in these samples

[43] For the lognormal distribution, any VF between 1 and 6.82 (the maximum) corresponds to two different levels of variation, one with a CV below 6.75 and one with a CV above 6.75. In this guidance, VFs for the lognormal distribution refer to the lower of the two possible levels of variation.

[44] Alternatively, an equivalent method is to use the variability factor to define a distribution for the ratios of individual unit concentrations to the sample/lot mean, then take samples from this distribution and multiply them with the sample/lot mean to obtain simulated unit concentrations.

may be less than the variance of the overall distribution. However, little or no evidence of such a negative correlation was observed in samples from market surveys by Hill and Reynolds (2002).

In principle, these complications regarding the variation of variability factors could be modelled statistically, but further research would be needed to develop and implement such an approach[45]. Therefore, the Panel recommends simpler models with fixed variability factors for **basic probabilistic assessments**. More sophisticated models with variable variability factors could be considered in the future as potential options for **refined probabilistic assessments**.

In some comparative calculations the Panel found that setting the variability factor for a lognormal distribution to 1 (i.e., no unit-to-unit variation) made no discernible difference to the resulting exposure distributions at the percentiles examined in that study (EFSA, 2007a). This was surprising because, logically, multiplying residues by a factor >1 must increase residues at some percentile. A preliminary investigation by the US EPA also suggested that the variability of unit-to-unit residues within a lot appears to have little impact on probabilistic modelling of the 99.9th percentile exposures (D. Miller, personal communication). In probabilistic modelling of cumulative exposures to triazoles, it was found that including variability factors had little effect when the lot mean residues were based on monitoring data, but a marked effect on higher percentile exposures when the sample mean was set to the MRL (van Klaveren et al., 2010). Furthermore, the variability factor in general would logically be expected to have an impact on the extreme tail of the exposure distribution, perhaps at extreme percentiles (e.g., above 99.9) that were outside the range of the earlier studies. Therefore, unit-to-unit variability is omitted from *optimistic model runs* for **basic probabilistic assessments**, but included in *pessimistic model runs* for commodities with a unit weight over 25g[46].

The appropriate form for the distribution of unit residues is uncertain. When the model for residue variability is sample-based, a distribution with an upper bound should be used (see below). When the model is lot-based, either distributions with or without upper bounds could be chosen, although it may be difficult to justify any specific choice of absolute upper bound. A simple choice of unbounded distribution for a lot-based model is the lognormal distribution. However, marked deviations from the lognormal distribution have been found in the marketplace, in some cases being multimodal, partly due to some lots in trade containing mixtures of units with different treatment histories (Hill, 2000). Of 116 datasets on unit residues from market surveys examined for EFSA (2005), the majority show marked deviations from lognormality, many being very strongly bimodal with a large proportion of non-detects that are clearly separated from the distribution of positive residues (P Craig, personal communication)[47].

In most of the probabilistic modelling conducted for another Opinion (EFSA, 2007a), the PPR Panel did not use a distribution of variability factors but instead set the variability factor for a lognormal distribution to a fixed value of 6.82. This was stated to be conservative, on the grounds that 6.82 is the maximum variability factor consistent with a lognormal distribution of unit values. In fact, as explained earlier, although 6.82 is the highest variability factor consistent with a lognormal distribution, there is no maximum for the variance of the lognormal, so there is no absolute worst case.

[45] For example, one area of potential research is using maximum likelihood techniques to investigate mixture distributions. One software tool that uses this method has been reviewed by the US EPA's Office of Pesticide Program's FIFRA Scientific Advisory Panel. Additional information is available at the US EPA's Scientific Advisory Panel website at: http://www.epa.gov/scipoly/sap/meetings/2000/february.

[46] It is proposed to use the same threshold of 25g for applying variability factors as JMPR (2003). If it appears possible that unit-to-unit variability of commodities with unit weight under 25g might have a significant influence on the outcome of an assessment, then this could be considered as part of the evaluation of unquantified uncertainties (section 8) and subsequently quantified if appropriate.

[47] Even when units share a common treatment history, they may not follow a lognormal distribution. Of 30 datasets of unit residues from supervised trials examined for EFSA (2005), 19 showed deviations from lognormality at $P<0.05$ and, of these, 10 at $P<0.001$ (Shapiro-Wilks test) (Peter Craig, personal communication). Deviations from the normal distribution (without taking logarithms) were much stronger.

When a sample-based conceptual model is considered, unit-to-unit variability is modelled as relating to samples taken from the marketplace rather than to lots in the marketplace (see section 4.2.2 above). In this case, there is an absolute worst case for the maximum unit residue in each sample which occurs when all of the measured residue for the sample derives from just one unit and the remaining units contain zero residues. This situation can be represented by a beta distribution, where the individual units in the sample can have residues between zero and the maximum with the constraint that the average must equal the mean for the sample. The worst case situation where all residue is contained in a single unit in each sample corresponds to a Bernoulli model.

For both the lognormal and beta models, assuming a high variance will overestimate the true proportions of high residues but underestimate the proportions of low residues. This creates an overall distribution which is conservative at high percentiles but unconservative at lower percentiles. Adjustments can be made to avoid underestimation at all percentiles, for example by resetting simulated values below the sample mean with the sample mean itself (van der Voet et al. 2003). However, this may result in large overestimation of lower tail residues which could cause large over-estimation of upper tail exposures when individual intakes are summed over multiple units of focal and background commodities.

For modelling unit-to-unit variability in *pessimistic model runs* in **basic probabilistic assessments**, the Panel seeks a simple model with a fixed variability factor or coefficient of variation which generates distributions of residues that reliably fall above the true distribution but not to an unrealistic or extreme degree. Early simulations by the Panel suggest that assuming a lognormal distribution with variability factor of 6.82 (as in EFSA 2007a) will generate an excessive proportion of very high residues. The Bernoulli model is clearly unrealistic in that it assumes a single unit contains all the residue in every sample. It is not possible to identify, a priori, what combination of model assumptions would meet the criteria indicated above. Therefore, further simulations are needed to evaluate the realism of distributions generated by different combinations of assumptions, as a basis for making final recommendations for regulatory assessment (see Conclusions section).

If sensitivity analysis shows that the assumptions made for unit-to-unit variability in the basic assessment have a significant impact on the risk management decision, then more sophisticated modelling of variability factors should be considered as an option for refined probabilistic assessment.

In some cases, measurements of residues in individual units may be available for the pesticide and commodity under assessment. In this situation, it may be attractive to use a variability factor derived from those measurements. However, it is not advisable to rely entirely on a single estimate of the variability because it will not reflect the known variation of variability factors (EFSA, 2005a). Furthermore, if the data derive from supervised trial conditions, this may underestimate unit-to-unit variation in the marketplace (EFSA, 2005a), especially when treated and untreated lots are mixed. If the choice of variability factor appears critical to the outcome of the assessment and risk management decision, consideration could be given to requiring multiple studies of the variability factor conducted under a realistic range of conditions (for further guidance see comment number 35 in EFSA (2006a)).

4.2.10. Simulating the combinations of residues encountered by consumers

When modelling dietary exposure in the presence of variation within and between lots or samples of a commodity, it is important to consider the way in which individual consumers select samples and units for consumption. A simple assumption might be that each unit consumed is selected at random from a different sample or lot, which is in turn selected at random from the distribution of samples or lots in the marketplace. In reality, however, a consumer who eats two units of the same commodity on the same day will often – but not always – take them from the same purchase, and therefore potentially from the same sample or lot. The effect of this behaviour on the exposure distribution depends on whether the model is based on samples or lots. In both cases, taking two units from the same lot will tend to increase the variance in exposures, as it will increase the proportion of cases in which both

units come from a high residue sample/lot, and the proportion in which both come from a low residue sample/lot. However, in the sample-based model, the sum of the two residues cannot exceed N times the sample mean, where N is the size of the sample, whereas in the lot-based model the two units would be drawn independently from the distribution for unit-to-unit variability, so higher exposures will be possible.

In reality, the true behaviour of consumers will lie somewhere between the extremes: sometimes multiple units of the same commodity will come from the same sample or lot, and sometimes from different samples or lots. The Panel considers that modelling this realistically using data or expert judgment is sufficiently complex to be reserved for **refined probabilistic assessments**. In **basic probabilistic assessments**, *optimistic model runs* should assume every unit is selected at random from a different sample or lot while, in *pessimistic model runs*, units of the same commodity consumed on the same person-day should be sampled at random from a single sample or lot, i.e., based on the same simulated sample or lot mean residue. If the model is sample-based, units from the same sample should be constrained not to exceed the maximum total residue.

4.2.11. Residue measurement uncertainty

Both the values below the LOR and the reported values are subject to measurement uncertainty. Measurement uncertainty is not normally considered in regulatory assessment in the EU, although allowance is made for it in criteria for confirming infringements of the MRL (SANCO, 2011a). Monitoring data available at the EU level has normally not been accompanied by information on measurement uncertainty[48]. It would be possible to include a generic estimate of measurement uncertainty based on the Horwitz equation, which relates the degree of variation to the level of the measured residue (Horwitz, 1982). However, the level of measurement uncertainty will generally be relatively limited, as most laboratories will comply with the SANCO requirement that analytical performance for routine recoveries should be within a Relative Standard Deviation of ±20% for within-laboratory reproducibility (SANCO, 2011a, paragraph 66) and an expanded uncertainty figure of ±50% in general covers the inter-laboratory variability between European laboratories (SANCO, 2011a, paragraph 91). A recent case study on pyraclostrobin exposure (M. Bakker, RIVM, personal communication) found that quantifying measurement uncertainty by a coefficient of variation (CV) of 50% led to over-estimation of the 99.9th percentile by 30%, and recommended applying a correction for this. However, at lower percentiles of exposure the impact was less, for example the 95th percentile was over-estimated by only 10%. In an example involving a single substance and commodity, EFSA (2007a, p. 70) found that sampling uncertainty had more impact than measurement uncertainty, although it is not known how general a result this would be. Overall, the Panel considers that measurement uncertainty is likely to have less impact on the assessment outcome than other sources of uncertainty, e.g., values below the LOR (which are set equal to the LOR in the pessimistic model but may be mostly true zeroes) and sampling uncertainty (which will usually be much larger than ±50%, as small samples are common). On this basis, the Panel considers it is not necessary to quantify measurement uncertainty, at least in the **basic probabilistic assessment**. Consideration could be given to including it in **refined assessments** if it was thought possible it would substantially change the assessment outcome (e.g., if the exposure estimates in the basic assessment are close to an acceptable limit and driven by a very small number of high residue values).

4.3. Processing factors

Processes such as cooking, peeling and juicing may decrease or increase the concentrations of residues in foods as eaten, compared to the levels in the raw agricultural commodities. For some foods, there may be more than one processing step between the commodity for which residue data are available and the food as eaten (e.g., wheat is milled to flour, then flour is baked in a pizza or other food).

[48] Such information may be provided more often in future, due to the requirements of ISO 17500.

Processing factors (defined as the ratios between concentrations in processed and unprocessed foods) are used in deterministic dietary exposure assessments to take account of these changes. The processing factor includes effects of processing on the chemical (e.g., degradation) as well as the food item (e.g., shrinkage). Effects on the food item are normally included in food conversion factors so when both factors are used it is important to avoid double-counting (see below).

Processing studies are defined in the relevant guidelines (EC 1997, OECD 2008). According to the current OECD Guidance Document (OECD 2008) in principle, for every crop having residues and being processed, a set of processing studies should be conducted (per pesticide). It should be possible to extrapolate the processing factor for the given pesticide to all crops within the given crop group undergoing the same procedure. Table 1 in OECD 2008 lists Processing Procedure Types and Recommended Extrapolations Using Typical RACs. Despite the detailed requirements, the extent to which the available studies represent the full range of processing practices is uncertain. Therefore, in **basic probabilistic assessments**, the Panel recommends using alternative assumptions to explore more and less conservative assumptions. *Optimistic model runs* should take the values used in deterministic assessment, including in the report a justification on which processing factors are used and which foods they are applied to. *Pessimistic model runs* should use a processing factor of 1 or the highest individual measured value available from processing studies, whichever is higher.

It is important to avoid double-counting of processing effects on the residue and the food item. When a processing factor is used for a food, as in the optimistic assessment, a conversion factor should not be used for the same food[49]. When the processing factor is not used (or set to 1), as in the pessimistic assessment, an appropriate conversion factor for that food should be used.

If sensitivity analysis shows that these alternative assumptions have a significant impact on the risk management decision, then more detailed modelling of processing factors should be considered as an option for **refined probabilistic assessment**. This refined modelling should include, for example, developing distributions to represent the variability of processing effects and the associated uncertainty based on available data and/or expert judgement. Potentially relevant data includes direct measurements of processing and also other information that may assist in making expert judgements, including physico-chemical characteristics of the residual compounds. Data on water content of foods (e.g., US EPA, 1996) may assist in estimating the concentration of residues when water content is reduced during cooking or drying if this is not covered by the residue conversion factor.

4.4. Residues in prepared foods

Much prepared food is purchased as such, but some is prepared from raw commodities at home. In some situations, an individual may purchase some raw commodity and consume some of it raw and some in prepared foods, all derived from a single lot or sample of raw commodity. For person-days where this occurs, there will be a positive correlation between residues in the raw and prepared foods which may contribute to the high percentile dietary exposures. Ideally, one would model these correlations using data or expert judgments on the proportion of each food prepared at home, but this would be complex to implement and is therefore only suitable for **refined probabilistic assessments**.

In *optimistic model runs* for **basic probabilistic assessments**, it may for simplicity be assumed that all prepared food is purchased ready-made. In *pessimistic model runs*, it should be assumed that all prepared food is prepared at home except for those prepared foods where this is not reasonable (as

[49] Where a processing factor exists for residues in the food as consumed, this should be used in the optimistic assessment together with the consumption of food as consumed, without conversion to raw commodity. When there is no processing factor for residues in the food as consumed, then the food as consumed should be converted to raw commodity and combined with residues measured in the raw commodity. In the latter case the residues in the food as consumed may be over-estimated, because allowance has been made for the change of commodity mass between the raw and processed food but not for any reduction in residues during processing.

listed in Table 1 in OECD Guideline 508). Where the same person-day includes more than one food from the same commodity, whether prepared or raw, they should all be assumed to come from the same sample or lot. This will occur sometimes though not frequently, e.g., a person who purchases oranges and consumes some raw and some after juicing. To represent this in the model, a single value should be drawn for the raw commodity (representing the mean of the single lot or sample for that consumer) and used as the starting point for sampling unit residues for portions of that commodity eaten both raw and as home-prepared food.

4.5. Residues in water

Exposure via drinking water should be included in the pessimistic model runs for basic probabilistic assessment. In the *pessimistic model run* for **basic probabilistic assessments**, the legal limit for residues may be used, i.e., 0.1 ppb for a single substance and 0.5 ppb for a cumulative assessments (comprising 0.1 ppb for each of the five most toxic substances in the cumulative assessment group). For the *optimistic model run* in a basic assessment, residues in via water could be assumed to be zero. Using monitoring data on concentrations in drinking water is challenging, due to regional and seasonal variations, but could be considered as an option for **refined probabilistic assessment**.

4.6. Combining consumption and residues by Monte Carlo simulation

When both consumption and residues are represented by distributions, it is necessary to combine these in a suitable way to estimate the resulting distribution of acute dietary exposures. This is generally done by Monte Carlo simulation, combining dietary records (person-days) with residue values sampled at random from the distributions representing variation in residues between and within lots.

In order to produce confidence intervals showing the uncertainty that has been quantified, a two-dimensional Monte Carlo (2D MC) procedure is used (e.g., Vose, 2008). This is illustrated diagrammatically in Figure 4.

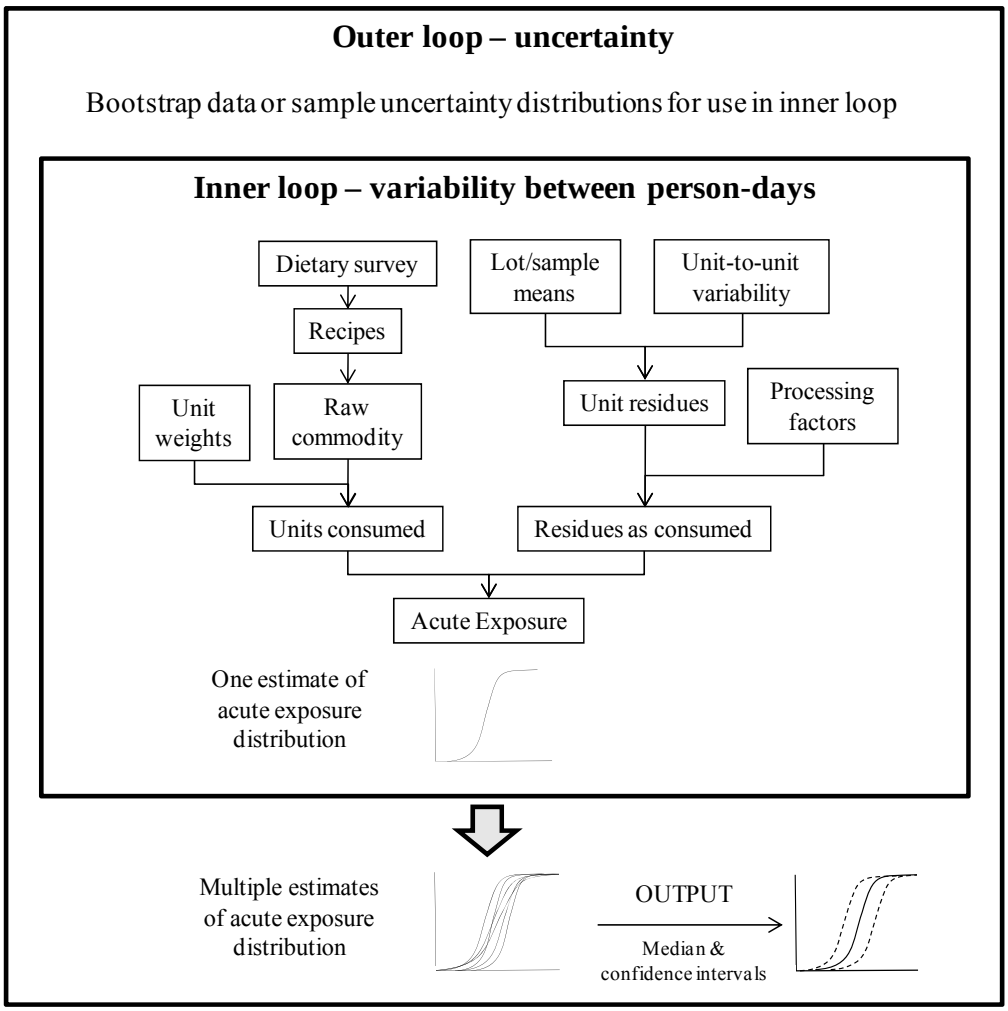


Figure 4: Illustration of procedure for two-dimensional Monte Carlo simulation of uncertainty and variability of acute dietary exposures. See text for details.

The 2D MC procedure comprises an inner and outer loop: the inner loop simulates variability of exposure between person-days, and the outer loop simulates those uncertainties that are being quantified in the assessment (Figure 4). The calculation of exposure is repeated many times in the inner loop, simulating different person-days by drawing different values from the data or distributions representing those parameters for which variation is being quantified. This constitutes one iteration of the outer loop, and generates one estimate of the distribution of exposures, as illustrated in the inner box in Figure 4. The outer loop is repeated many times, each time taking different distributions or resampled datasets for the parameters for which uncertainty is being quantified. Each outer loop generates one estimate of the distribution of exposures, which can be plotted together and used to derive a median estimate and confidence intervals for the distribution of exposures, as illustrated diagrammatically at the bottom of Figure 4.

Variability and uncertainty are simulated in two ways: bootstrapping, where random samples are drawn from input data, or statistical modelling, where random samples are drawn from distributions derived from a statistical model of the variability and uncertainty for an input of the exposure assessment. In the approaches recommended by the Panel for **basic probabilistic assessment** of acute exposure, bootstrapping is used for consumption data and statistical modelling is used for the distribution of mean residues between lots or samples. In **refined probabilistic assessment**,

variability and uncertainty may be quantified for additional inputs (e.g., variability factors, processing factors, unit weights). As the underlying datasets for these inputs will generally be small, their variability and uncertainty should preferably be quantified using statistical models as empirical bootstrapping will only generate values contained in the underlying data. Also, in refined assessment, consumption may be modelled parametrically rather than by empirical bootstrapping (see section 4.1). Table 2 (see earlier) identifies which variables and uncertainties are modelled empirically, parametrically and with fixed values in basic assessments.

In **basic probabilistic assessments of acute exposure**, sampling uncertainty for consumption should be quantified in the outer loop by resampling the person-day records in the raw survey data. In each outer loop, a set of records equal to the number in the original survey (*S*) is drawn by random sampling with replacement. This generates a different sample of *S* consumption records for use in each inner loop.

In *optimistic model runs*, empirical bootstrapping is also used to quantify sampling uncertainty for lot/sample residues following the same procedure as for consumption. In each outer loop, a number of residues *R* equal to the number in the observed data is drawn at random with replacement from the observed data. In *pessimistic model runs*, uncertainty and variability for lot/sample residues is modelled parametrically, using the equations (1)-(3) in sections 4.2.5 and 4.2.7. Each outer loop samples one value for the proportion of positive residues, defining a single binomial distribution for use in one inner loop. Each outer loop also samples one mean and combines this with a standard deviation (derived from the variability factor) to define a lognormal distribution which is then used in the inner loop to represent the variation of positive residues. Individual residues are then simulated in the inner loop by sampling from those distributions.

In order for the inner loop to adequately explore the variability in exposure that results from the many possible combinations of consumption and residues, it is necessary in the inner loop to simulate a much larger population (*P*) of person-days than in the original survey. This is done by drawing the required number of records *P* at random with replacement from the *S* records selected in the outer loop[50].

Sufficient inner and outer loops should be simulated to produce stable results which do not change materially if the assessment is repeated. The numbers required may vary from assessment to assessment. Boon and Van Klaveren (2003) recommended simulating 100,000 person-days, but larger simulations may be needed, especially where rarely-eaten foods are important and/or when high percentiles of the exposure distribution approach or exceed toxic reference values. Therefore the Panel recommends that, whatever number of inner and outer loops is simulated initially, the assessment should be repeated at least 3 times[51]: if the output distributions and their confidence intervals are similar and they do not alter the implications for risk management, then the degree of replication may be considered sufficient. Otherwise, the number of inner and outer loops should be increased until repeated runs give stable outputs.

It is important to keep in mind the limitations of methods used to quantify variability and uncertainty. Because bootstrapping is limited to the range of observed values, consumption data contributing to the upper tail of the estimated exposure distribution should be examined (see drill down in section 7), and the potential for higher consumption of those foods to occur in reality should be considered. This is especially important for less-commonly consumed foods which may be represented by only a very small number of records in the data. Parametric modelling of residues using the lognormal distribution

[50] In some models, the required number of person-days has been obtained by replicating the consumption data rather than resampling it. The results should be similar, provided the number of person-days simulated is large enough (judged by stability of the outputs as described in the text).

[51] With different sequences of random samples (i.e., different random number seeds). Three repetitions are recommended because 3 similar results provide more assurance (compared to 2 repetitions) that the number of iterations is sufficient to produce stable results.

will sometimes generate impossibly high residues. Therefore, simulated residues contributing to the upper tail of the estimated exposure distribution should also be examined to evaluate their realism (see section 7). Finally, many potential sources of variability and uncertainty are not quantified in the proposed approaches for basic probabilistic assessment, so it is essential to evaluate these subjectively in every assessment (see section 8).

5. Modelling chronic exposure

Chronic dietary exposures should be estimated for relevant scenarios, identified as indicated in section 2. Chronic exposure is averaged over time but varies between individuals due to differences in their dietary habits.

The basic inputs required for modelling chronic dietary exposure are the same as those for acute exposure: the amounts of pesticide residue that are present in and on foods and the types and amounts of those food consumed in a person's diet. However, to model chronic exposure it is necessary to have consumption data for multiple days per person, unless estimates of within-person, between-day variation are available from another source. Again, a number of additional variables are also used. Some of these are adjustments required to allow the assessment to be conducted with the types of data that are normally available, while others allow the user to take account of factors that may modify exposure. They include:

- food conversion factors, to convert composite food products as recorded in dietary surveys (i.e., as eaten) to their individual ingredients;
- processing factors, to take account of changes in nature and amount of residues during the processing of raw agricultural commodities into processed commodities or ingredients;
- percentage of the commodity that is treated with the pesticide under assessment, to take account that this is generally less than 100%.

Variability factors and unit weights are not needed for chronic assessments. This is because short-term variability of residues within and between lots or samples is not relevant when modelling chronic exposure. Instead, chronic exposure is estimated using the average residue for each commodity. In this respect, modelling chronic exposure is simpler than acute exposure. However, modelling long-term average exposures using consumption data from short-term dietary surveys introduces significant challenges in chronic exposure assessment.

The following sections discuss the possibilities for how each element of the chronic exposure model could be handled in a probabilistic assessment, and the difficulties that arise. They also explain the Panel's conclusions on which options should be used in optimistic and pessimistic runs for a basic probabilistic assessment, and which of them might be options for refined assessment. These conclusions are summarised in Table 3.

It is important to note the different nature of the optimistic model run in chronic assessment, when compared to acute assessment. In the acute assessment, the optimistic run is expected to underestimate exposure, so if it raises concern it is unlikely this can be removed by refinement and it may be advisable to stop the assessment (see Figure 1 in section 2). Whereas in chronic assessment, the optimistic model run may overestimate exposure (although to a lesser extent than the pessimistic run), and its role is to indicate whether refinement should include parametric modelling of intakes rather than the simple Observed Individual Means (OIM) approach (see 5.1.2 for more detail).

Table 3: Summary of recommended approaches for chronic dietary exposure assessment (see the indicated text sections for detail). Each aspect applies to all types of assessment scenario (see Table 1). The Panel expects that many of the approaches for basic assessment are likely to become available to users as built-in options in probabilistic software. Key to probabilistic methods in basic assessment (not shown for refined assessment): * fixed value; § variability modelled empirically, §§ variability modelled parametrically, ¶ uncertainty modelled by bootstrapping, ¶¶ uncertainty modelled parametrically.

Assessment component	Basic assessment		Options for refined assessment include:	Section no.
	Optimistic	Pessimistic		
Modelling consumption (modelled as exposure, after combination with residues)	Observed Individual Means method + bootstrap §¶	Observed Individual Means method + bootstrap; examine which commodities contribute to upper tail exposures. §¶	Parametric models	5.1.2
Water consumption	Zero	Treat as food if included in dietary survey, or use deterministic estimate from drinking water assessment.*	More sophisticated estimates	5.1.4
Food conversion factors (recipes)	Use available recipe databases*		Quantify variability and uncertainty for foods driving exposure	5.1.3
Residue definitions	Use residue definition for risk assessment, applying conversion factor where appropriate*. Evaluation of unquantified uncertainties.		Consider more sophisticated methods (see 5.2.1)	5.2.1
Unmeasured residues in animal commodities with authorised uses	Zero*	MRL or default MRL*	More sophisticated estimates (see 5.2.1)	5.2.1
Residue measurement uncertainty	Not modelled.	Not modelled.	Consider including if thought potentially important.	5.2.6
Mean residue for each commodity	Mean of available data*		Parametric models	5.2.2
Sampling uncertainty for mean residues	Empirical bootstrap of available data ¶	Empirical bootstrap of available data ¶		5.2.2
Treatment of residues below LOR (Level of Reporting)	Treat as true zeroes*	Set <LOR to LOR*		5.2.2, 5.2.4
Percent crop treated (when using supervised trials/feeding studies data)	Approximate estimate of % crop treated*	Assume 100% of crop treated*	Refined estimate of % crop treated and the uncertainty of this	5.2.4
Limited amounts of monitoring data	Use available data empirically	Use appropriate data from other countries, other commodities or supervised trials/feeding studies	Future options might include extrapolation between substances	5.2.3
Residues from animal feeding studies	Use measured value from 1x dose of feeding study, or set to zero*	Use measured value from 1x dose of feeding study, or set to MRL or default MRL (0.01)*	More sophisticated estimates (see 5.2.1)	5.2.1
No supervised trials data (as substitute for monitoring data)	If no monitoring or trials data, assume no residues.	Use appropriate trials data from other commodities	Future options might include extrapolation between substances	5.2.3
Residues for non-authorised use	Treat as for authorised uses	Treat as for authorised uses except set <LOR to zero*	Treat as for authorised uses	5.2.4

Processing factors	Value used in deterministic assessment.*	Distribution of estimates for mean processing factor, obtained by bootstrapping measured values. *	Quantify uncertainty using data and/or expert judgment	5.3
Residues in water	Zero	Assume legal limit (0.1ppb for single substance, 0.5ppm for cumulative assessment)*	Consider using monitoring data from water	5.4
Cumulative assessment	See section 6			6
Unquantified uncertainties	Optionally, evaluate using uncertainty table	Evaluate using uncertainty table	More sophisticated or quantitative evaluation	8

5.1. Consumption

5.1.1. Data organisation and adjustment

Although average (habitual) consumption over longer time periods is relevant for assessing chronic dietary exposure, available consumption data are from short term surveys. As for acute assessments, survey data reporting foods as eaten need to be converted to the appropriate quantities of raw commodities and expressed relative to body weight for use in chronic exposure modelling (see 4.1.1 and 4.1.3 for more details).

Rarely eaten foods may be under- or over-represented in the consumption surveys. Therefore, assessors should always check if the data include all commodities with approvals or positive residues. If not, then the assessor should consider their potential contribution as part of the evaluation of uncertainties and consider modelling them as an option for refined assessment (see 4.1.1).

5.1.2. Empirical use of consumption data and parametric modelling of chronic exposure

In chronic dietary exposure assessment, it is necessary to estimate long-term exposure using consumption data from dietary surveys that often cover only 2-4 days. This extrapolation may be done very simply in an empirical approach, referred to as the Observed Individual Means approach (OIM) (termed the naïve approach in Kipnis et al. 2009). This uses the observed mean consumption over the recorded days for each individual to calculate mean exposures and then treats these as estimates of long term exposures. Alternatively, the extrapolation may be done parametrically by fitting a statistical model that separates within- and between-individual variation in consumption or exposure and uses this to estimate long-term average exposures.

Due to the limited duration of dietary surveys, the OIM approach tends to show exaggerated differences between individuals due to short-term variations in diet over time that tend to average out over longer time periods. For example, if an individual happens to purchase a kilogram of pears in the survey period, he is more likely to consume pears on each survey day and his average consumption of pears in the survey may greatly overestimate his real long-term average. Therefore, the OIM approach will tend to over-estimate upper tail exposures in chronic assessments because a short-term survey is likely to over-represent the frequency of individuals consuming a food every day, although underestimation of upper tail exposures may also occur[52].

[52] For symmetrical distributions the 50th percentile (median) is the turning point between under-estimation and over-estimation. However, for right-skewed distributions (as are typical for pesticide exposure), the turning point is a higher percentile. Moreover, this percentile increases when the short-term variation gets larger relative to the between-person variation, as demonstrated in a simulation study by Goedhart et al. (2012). It may therefore occur that percentiles of interest for decision-making are under-estimated rather than over-estimated by the OIM method.

The OIM approach is liable to underestimate the proportion of the population that is ever exposed. This is because only a proportion of the persons who eat a commodity will happen to eat it during the short period when their diet is surveyed.

Parametric approaches to estimating long-term food consumption are intended to overcome the limitations of the empirical approach. They avoid the potential biases identified above, and can estimate the frequencies of high consumption events that were not observed in the dietary survey. However, in order to produce reliable estimates of intake, it is essential to take account of the complex correlations that occur between consumption of different foods. A simple option is to combine the observed consumption data for each food type with the mean residues for the same food to obtain estimated daily exposures, and then apply a parametric model to those exposures. This incorporates the correlations present in the data and avoids the need to model them explicitly, although they are subject to sampling uncertainty, especially for correlations between less-commonly consumed foods.

Researchers have developed in recent years a range of statistical approaches to model dietary patterns, including accounting for correlations between foods and also the dependence of consumption patterns on other variables (covariates) including age and body weight. Examples include the approaches of Nusser (1996, 1997), Slob (1993, 2006), Tooze et al. (2006), Allcroft et al. (2007), de Boer et al. (2009), Kipnis et al. (2009), Kennedy (2010) and Zhang et al. (2011). Most of these methods are similar in that they allow for the presence of zero exposures in addition to positive values by having two parts, one for modelling frequencies and one for modelling amounts. The approaches differ in their assumptions about the distributions describing variation in frequencies and amounts of consumption or exposure and in the degree to which they account for the effects of covariates.

The relative suitability of these varied approaches for statistical modelling of chronic exposure is a subject of current research. De Boer et al. (2009) compared the beta-binomial normal (BBN) method with the Iowa State University Foods (ISUF) model. They found that neither model is suitable for use when the distribution of exposures is multimodal. Although the ISUF model includes a spline transformation that will always give a normal distribution, this transformation is not compatible with the assumption that between and within consumer variances are additive (de Boer et al, 2009, page 1448). When a logarithmic or power transformation results in an approximately normal distribution of exposures, de Boer et al. (2009) prefer the BBN model over the more complex ISUF model. De Boer et al. (2009) conclude that the choice of appropriate models should be made on a case-by-case basis and that more research is necessary to develop a method that is applicable to multimodal exposure distributions. In a more recent simulation study by Goedhart et al. (2012) as part of the ETUI project[53], several methods were compared in example cases where the exposure was non-daily, as is true for pesticides. Amongst these methods, very similar results for upper-tail exposure estimates were obtained for parametric methods based on the logistic-normal normal (LNN) or NCI method (Tooze et al. 2006) and the beta-binomial normal (BBN) model (Slob 2006, de Boer et al. 2009). When strong correlations existed between the frequencies of consumption and the amounts of consumption, the use of LNN was preferable to BBN. The results of all these methods were usually much closer to the true simulated exposures than those obtained with the OIM method, which were either far too high (in the extreme upper tail) or far too low (in most of the rest of the distribution).

The parametric models discussed above assume that, after some suitable transformation, exposure amounts will follow a normal distribution. However, in scenarios for authorization, approval or MRL setting (see section 3 and Figure 2), residues in the the focal commodity are based on field trial data, which typically comprise much higher values than monitoring data. As a consequence there is a strong prior expectation of severe bimodality for the exposure distribution. Most parametric models will not

[53] The report of a workshop conducted under the EFSA-funded project 'European Tool Usual Intake' may be found at: http://www.efsa.europa.eu/en/supporting/pub/86e.htm (van der Voet and van Klaveren 2010). The final reports of the project may be found at http://www.efsa.europa.eu/en/supporting/pub/300e.htm (van Klaveren et al. 2012) and http://www.efsa.europa.eu/en/supporting/pub/299e.htm (Goedhart et al. 2012).

fit multimodal distributions well. The most likely consequence is that a very broad normal distribution will be fitted to the true multimodal pattern, leading to a large over-estimation of the upper tail. A possible approach to this problem could be to use parametric models for the focal commodity and background separately, and then add the resulting distributions (Model-then-add approach, see van Klaveren et al. 2012).

Taking account of the limitations of the existing approaches and the significant statistical expertise required to use them correctly, the Panel recommends that **basic probabilistic chronic exposure assessments** should use the OIM approach for both *optimistic* and *pessimistic* assessments. Both *optimistic* and *pessimistic model runs* should include bootstrapping of the dietary records, to indicate the degree of sampling uncertainty affecting exposure estimates. In addition, it is essential to check the dietary records that generate the upper tail exposures for *pessimistic model runs* to identify which foods contribute significantly to the exposure, and consider whether they might be consumed more frequently by some consumers than was found in the dietary survey. If so, the OIM approach may underestimate the upper tail residues, which should then be investigated by parametric modelling as part of a refined assessment.

It is important to emphasise that the OIM approach tends to overestimate upper-tail exposures, as explained earlier in this section, and therefore changes the nature of the *optimistic model run* in the case of chronic assessment. Although the treatment of residues in the optimistic run will tend to underestimate exposure (see following sections), this may not be sufficient to counteract the tendency to overestimation caused by the OIM approach. This must be taken into account when interpreting the results of the optimistic and pessimistic model runs for chronic assessment. Specifically, if one or both model runs generate exposures that raise concern, it is possible that both are overestimates and that refinement using a parametric approach in place of the OIM may remove the concern[54]. If the optimistic run does not raise concern but the pessimistic run does raise concern, this implies that refinement of the treatment of residues (e.g., non-detects, processing factors, animal commodities) may be sufficient to remove the concern without the need to move from the OIM approach to parametric modelling. Thus even though the 'optimistic' model run for chronic assessment is not literally optimistic, it is useful in helping the user decide between the options for refinement.

Parametric modelling of exposures may be considered as an option for **refined probabilistic assessment** of chronic exposure. In cases where the distribution of exposures is approximately normal (after logarithmic or power transformations if needed), and there is no evidence of correlation between the frequencies and daily amounts of consumption for each food[55], application of the LNN or similar model may be appropriate. If these conditions are not met, other parametric solutions may be considered case-by-case, with the aid of expert statistical advice.

These recommendations supersede the view expressed by the Panel in its opinion on cumulative risk assessment where it stated that using the empirical (OIM) approach for chronic assessments is "entirely inappropriate" (EFSA, 2008, footnote 19 on page 35). The primary reason for the previous view was to avoid the tendency of empirical approach to over-estimate the upper tail of the exposure distribution. The Panel's new recommendation recognises that adequate parametric approaches are not yet developed for all situations (especially for multimodal distributions) and require a high level of expertise, while the simplicity and likely conservativism of the empirical approach make it suitable for a basic pessimistic probabilistic assessment. However, the potential for including other parametric approaches in the basic probabilistic assessment should be kept under review as new developments emerge.

[54] This contrasts with acute assessment, where the optimistic model is expected to underestimate exposure and therefore, if it raises concern, it is unlikely that this can be removed by any refinement option (see section 2).

[55] This can be examined by using box plots to compare daily consumption amounts for subsets of individuals who consumed the food in question on different numbers of survey days (0, 1, 2, etc.).

In general, it is essential that users be aware of the limitations of approaches used for modelling consumption and take them into account as part of the evaluation of unquantified uncertainties (section 8).

5.1.3. Food conversion factors

Mean food conversion factors should be used in chronic dietary exposure assessments. In **basic probabilistic assessments**, the Panel recommends to use the same values for these parameters as are used in deterministic assessments (see section 4.1.3). **Refined probabilistic assessments** could use means estimated from additional data, when available from appropriate and relevant studies. Unit weights are not required for chronic assessments.

5.1.4. Water consumption

Water consumption should be handled for chronic assessments in the same way as for acute assessments (see section 4.1.4).

5.2. Residues

5.2.1. Data organisation and adjustment

As previously mentioned, mean residues are relevant for modelling chronic dietary exposure, and there is no need for modelling between-lot, between-sample or between-unit variation as in acute assessments. However, the individual residue values underlying the means are required as input for modelling to enable bootstrapping to quantify the impact of sampling uncertainty on the mean values.

All considerations regarding organisation and adjustment of residue data for modelling acute exposure (section 4.2.1) apply similarly for modelling chronic exposure. The exception to this is that, for animal commodities, the measured value from the 1x dose in a feeding study should be used as the estimate for the mean residue in that commodity, without generating additional estimated values (c.f. section 4.2.1).

5.2.2. Modelling of residues

Modelling of variation in residues within and between lots or samples of a commodity is not needed for chronic dietary exposure assessment which should be based on the mean residue level for each commodity taking into account data both above and below the LOR. In both *optimistic* and *pessimistic model runs* for **basic probabilistic assessment** of chronic exposures, empirical bootstrapping should be used to give an indication of the sampling uncertainty of the mean values. The quantification of sampling uncertainty will be approximate, especially for smaller datasets. However, this problem is less severe than for acute assessment because bootstrapping performs better for the mean than for distribution tails and because sample size is increased by inclusion of data both above and below the LOR. Parametric modelling of the uncertainty of mean residues could be considered as an option for **refined probabilistic assessment**.

5.2.3. Using residue data from different sources to increase sample size

Although mean residues are less influenced than distribution tails by sampling uncertainty due to limited data, it is still desirable to reduce the uncertainty by combining data from different sources where appropriate. This is referred to as extrapolation.

In *optimistic model runs* for **basic probabilistic assessments**, no extrapolation is necessary. The assessment may be conducted using only the residue data that are available. This will tend to

underestimate dietary exposure because it will assume that residues are absent in commodities that have no measured values[56].

In *pessimistic model runs* for **basic probabilistic assessments** of chronic exposure, all available data that are appropriate for extrapolation should be used to estimate the mean residue for every commodity that has an authorised use or MRL for the substance under assessment, to take account of their potential contributions to exposure. The Panel recommends the following procedures:

- For commodities where monitoring data are appropriate (see Table 1 in section 3), use all relevant monitoring data for that commodity, including data from other countries where appropriate. If there are no monitoring data for the commodity in question, combine all relevant monitoring data from other commodities for which extrapolation is accepted. If there are no relevant monitoring data at all, either for the commodity in question or any other commodity from which extrapolation is appropriate, use data from supervised trials/feeding studies and proceed as in the following bullet.

- For commodities where supervised trial data or feeding studies are appropriate (see Table 1), use all available supervised trials for that commodity which are relevant to the GAP for the use under assessment. If there are no supervised trial/feeding study data for the commodity in question, substitute supervised trials or feeding studies from other commodities for which extrapolation is accepted. Again, use all the supervised trials/study feeding studies that are relevant. If there are no relevant data at all, substitute the MRL for the commodity in question.

Before using either monitoring or supervised trials data from other countries or commodities as described above, their relevance to the population and commodity in question should be critically assessed. Extrapolation should only be considered for pairs of commodities listed in guidance document SANCO 7525/VI/95, and only when it can be reasonably expected that the use and usage practices of the pesticide in question are the same in both commodities. Similarly, extrapolation between EU countries should follow guidance provided by SANCO (2011b). Extrapolation from countries outside the EU should only be considered when it can be reasonably expected that residues will be similar in both countries. All extrapolation should be fully documented and justified in the assessment report.

In a **refined probabilistic assessment**, extrapolation or combining of data from different sources may be done using appropriate statistical methods which quantify the associated uncertainty.

5.2.4. Handling of untreated commodity and residues below the limit of reporting

Monitoring data based on composite samples frequently contain a high proportion of values below the Limit Of Reporting (LOR), some of which may represent untreated commodity. Data from supervised trials all relate to treated commodity, but may also contain values below the LOR.

Statistical methods for dealing with data below the LOR were evaluated in a recent study by EFSA (2010b). As explained in section 4.2.7, pesticide residue datasets will often fail to meet the requirements proposed by EFSA (2010b), even when data for different commodities are pooled or additional data are collected. Therefore, in the case of pesticides, the PPR Panel proposes an alternative strategy, using different assumptions in the optimistic and pessimistic model runs to take account of the uncertainty in a way that is practical for basic probabilistic assessments.

For *optimistic model runs* in **basic probabilistic assessments**, the following procedures should be used:

- Where monitoring data are used, values below the LOR should be treated as true zeroes.

[56] For commodities with measured values, the sampling uncertainty of the mean residue will be higher when extrapolation is excluded, and may result in over- or underestimation, though this should tend to average out over multiple commodities.

- Where supervised trial data are used, values below the LOR may be treated as true zeros, as an optimistic assumption. In addition, allowance may be made for the proportion of each commodity that is expected to be untreated, by adding the appropriate proportion of zero values. In the optimistic basic assessment, the proportion of each commodity that is untreated can be an approximate judgment. Similarly, an appropriate proportion of the animal commodities can be considered as containing no residues as a result of consumption of untreated feedingstuffs by livestock. This may be represented by adding an appropriate number of zero values to the value(s) taken from the feeding studies as described in section 5.2.1.

For *pessimistic model runs* in **basic probabilistic assessments**, values below the LOR should be replaced with the LOR, as a conservative assumption, before bootstrapping. This applies to both monitoring data and supervised trials/feeding studies data. Where supervised trials/feeding studies data are used, it should be assumed that 100% of the commodity will be treated, i.e., no zero values should be added.

In both the optimistic and pessimistic model runs, if any values below the LOR relate to samples from regions of origin where there is no registered use for the pesticide and commodity in question, then those values may be considered as true zeroes. However, any positive concentrations recorded from regions without registered uses may result from illegal use and should be retained.

If sensitivity analysis shows that the treatment of values below the LOR has a large influence on estimated dietary exposures, consideration could be given to more sophisticated approaches in **refined probabilistic assessments**. This might include the use of advanced modelling approaches (e.g., those discussed by EFSA, 2010b) and more rigorous estimates of the proportion of each commodity that is or will be treated (see section 4.2.7 for discussion of this).

Note that using monitoring data for modelling implies an assumption that current or future levels of pesticide use are similar to those during the period to which the monitoring data relate which might cause either over- or underestimation of exposure. Similarly, using supervised trials/feeding studies data requires explicit assumptions about pesticide usage. Uncertainty about these assumptions should be quantified using formal expert elicitation or considered as part of the evaluation of unquantified uncertainties (section 9).

5.2.5. Addressing targeted sampling

Considerations regarding targeted sampling as described for modelling acute dietary exposure apply similarly here (see section 4.2.8).

5.2.6. Residue measurement uncertainty

In section 4.2.11, the Panel briefly reviewed the extent of measurement uncertainty in pesticide residue data and concluded it would have limited impact on assessment of acute exposure. Its impact on chronic exposure assessment is likely to be smaller still, because chronic assessment is based on mean residue levels which will be less affected by measurement uncertainty due to averaging over multiple samples. Accordingly, it is the Panel's view that measurement uncertainty may be ignored in **basic probabilistic assessment**, and considered in **refined probabilistic assessments** only in cases where the assessor considers it might substantially change the assessment outcome (e.g., if the exposure estimates in the basic assessment are close to an acceptable limit and driven by a commodity with a very small residue dataset).

5.3. Processing factors

Chronic dietary exposure should be estimated using mean values for processing factors, although when the mean value is based on more than one data value the individual values will be required as input for bootstrapping to examine the sampling uncertainty of the mean.

In **basic probabilistic assessments**, the Panel recommends using alternative assumptions to explore more and less conservative assumptions, as in acute exposure assessment. *Optimistic model runs* should take the mean values used in deterministic assessment. *Pessimistic model runs* should use a distribution of mean values obtained by bootstrapping the individual values used for each pesticide/commodity combination in deterministic assessment, as an approximate representation of the uncertainty due to the limited number of measurements. If sensitivity analysis shows that these alternative assumptions have a significant impact on the risk management decision, then more detailed modelling of mean processing factors should be considered as an option for **refined probabilistic assessment** (see section 4.3 for further discussion).

5.4. Residues in water

Residues in water should be handled for chronic assessments in the same way as for acute assessments (see section 4.5).

5.5. Simulation of chronic exposures

Chronic dietary exposure assessments generally assume that, for any given food type, the whole population is exposed to the same mean concentration over the long term. Therefore, the population distribution of exposure can be estimated without the need for probabilistic methods by simply combining each individual's consumption with the mean concentration for each food[57].

However, probabilistic methods are required to take account of uncertainty regarding either the consumption or concentration data. In both *optimistic* and *pessimistic model runs* for **basic probabilistic assessments** of chronic exposure, the Panel recommends this is done by bootstrapping the observed data. The procedure is illustrated in Figure 5. This is similar to the 2-dimensional Monte Carlo procedure used for acute exposure assessment (section 4.5 and Figure 4). However, for the basic probabilistic assessment, no sampling is required in the inner 'loop'. Instead, a single estimate of the mean concentration in each food is combined with the consumption data for each individual. In the outer loop, bootstrapping is used to quantify uncertainty for both consumption and residues. Uncertainty for consumption is quantified by resampling the person records in the raw survey data, keeping multiple days for each person together. In each outer loop, a set of records equal to the number in the original survey (S) is drawn by random sampling with replacement. This generates a different sample of S consumption records for use in each inner loop. Similarly, uncertainty in mean residues is quantified by resampling the residue data. In each outer loop, for each food, a number of residues R equal to the number in the observed data is drawn at random with replacement from the observed data, before calculating the mean.

For the basic assessment, the size of the inner loop is equal to the number of person records in the consumption data. Sufficient outer loops should be simulated to produce stable results that do not change materially if the assessment is repeated. The numbers required may vary from assessment to assessment. Therefore the Panel recommends that, whatever number of inner and outer loops is simulated initially, the assessment should be repeated at least 3 times[58]: if the output distributions and their confidence intervals are similar and they do not alter the implications for risk management, then the degree of replication may be considered sufficient. Otherwise, the number of outer loops should be increased until repeated runs give stable outputs.

Table 3 (see earlier) identifies which variables and uncertainties were modelled empirically, parametrically and with fixed values in the basic probabilistic assessment.

In **refined probabilistic assessments**, parametric models may be used to quantify uncertainties for one or more elements of the assessment (consumption, residues, processing factors, etc.), as indicated

[57] This is referred to as the 'individual-based deterministic approach' by EFSA (2010).
[58] With different sequences of random samples (i.e., different random number seeds).

in earlier sections. If consumption is modelled parametrically, the number of individuals simulated in the inner loop should be large enough to obtain stable estimates and confidence intervals at the levels of exposure that are of interest to risk managers.

It is important to keep in mind the limitations of methods used to quantify variability and uncertainty. Because bootstrapping is limited to the range of observed values, consumption data contributing to the upper tail of the estimated exposure distribution should be examined (see drill down in section 7), and the potential for higher or more frequent consumption of those foods to occur in reality should be considered. This is especially important for less-commonly consumed foods which may be represented by only a very small number of records in the data. Similarly, the uncertainty of mean concentrations for foods with small residue datasets may be poorly represented by bootstrapping. These and other unquantified sources of variability and uncertainty affecting the assessment should be evaluated subjectively (see section 8).

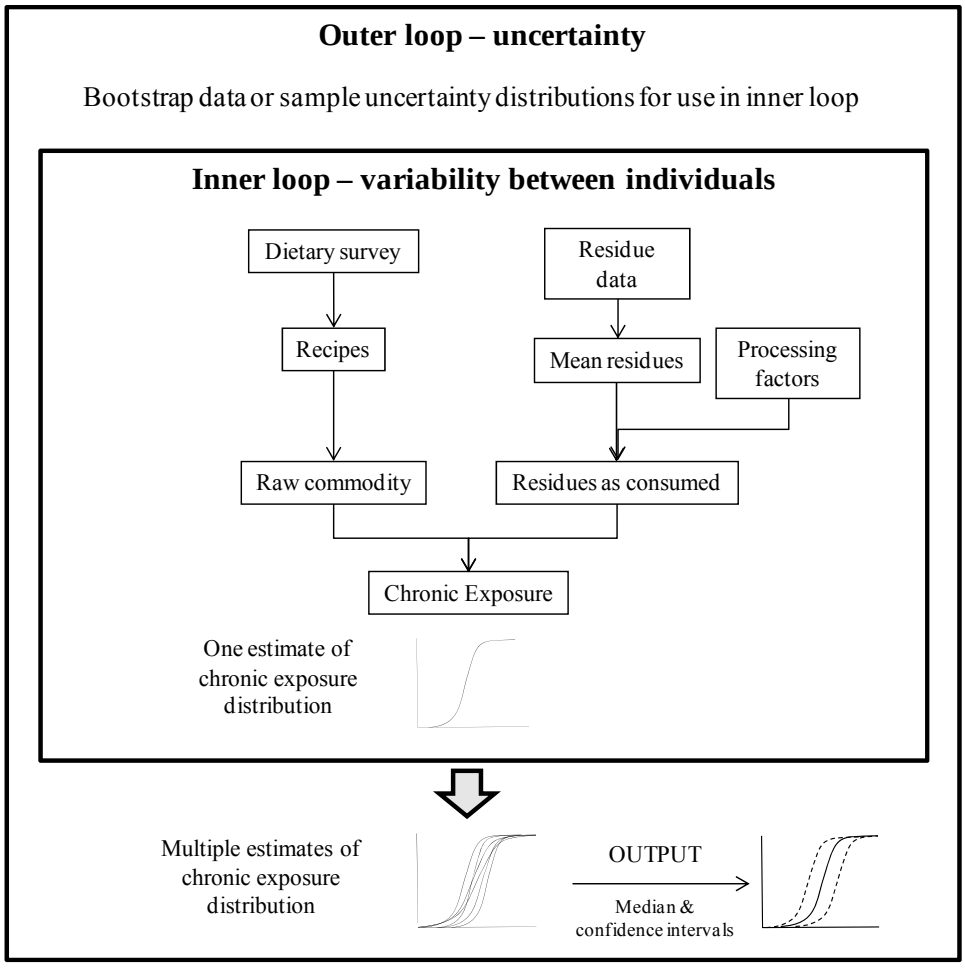


Figure 5: Illustration of procedure for simulation of uncertainty and variability of chronic dietary exposures. See text for details.

6. Additional methodology for cumulative exposure assessment

Cumulative assessments address the overall risk deriving from combined exposure to multiple compounds that share the same mode of action or that have similar effects but by different modes of action (EFSA, 2008). The Panel has previously reviewed approaches for cumulative assessments (EFSA, 2008) and evaluated their practical application to example substances (EFSA, 2009a).

The Panel has identified two aspects of cumulative assessment that impact on the methodology for probabilistic modelling: the methodology for cumulation of toxicity, and the need to quantify co-occurrence of residues for different substances in the same foods. These are addressed in the following sections.

6.1. Cumulation of toxicity

Methodology for cumulation of toxicity is relevant to this guidance on dietary exposure assessment because cumulative risk is assessed by combining the exposures of different compounds expressed as functions of their toxicities.

A basic consideration in cumulative risk assessment is the identification of the Cumulative Assessment Group (CAG), defined by EFSA (2008) as a group of chemicals that could plausibly act by a common mode of action, not all of which will necessarily do so. Methodology for assessing membership of the CAG will be addressed by the Panel in a separate opinion .

EFSA (2008) described the methods by which toxicity from exposures to different substances in the same CAG can be combined in a cumulative assessment. In order of increasing complexity, this can be by using a Hazard Index (HI) or adjusted Hazard Index (aHI), a Reference Points Index (RfPI), Relative Potency Factors (RPF), or physiologically based toxicokinetic and toxicodynamic modelling (PBTK and PBTD) approaches.

The HI and aHI are sums of the ratios of the individual compound exposures to their respective toxicological reference values. In the case of the aHI, any reference values that are not relevant for the specific common toxic effect upon which the CAG is based are replaced by effect-specific reference values. The RfPI approach sums the exposures to each pesticide expressed as a fraction of their individual reference points for the relevant effect (e.g., NOAEL or BMD10). The use of the RPF method requires identification of a reference pesticide for the CAG, i.e., an index compound (IC), and the relative potencies of the remaining compounds to the IC. Exposures are then summed as IC equivalents. Two other approaches, the combined margin of exposure (MOE) and the cumulative risk index (CRI), are reciprocally related to the RfPI and HI, respectively (EFSA, 2008).

In the case of approaches based on RfPs, uncertainty factors to account for extrapolation in toxicology data are applied to the end result. In approaches based on reference values, the toxicological uncertainty factors are applied to the individual compounds.

As discussed by EFSA (2008), the method chosen for cumulating hazard makes little difference when comparable toxicology studies and the same uncertainty factors are used.

The Panel's Opinion (EFSA, 2009a) examining the practical possibilities for assessing cumulative effects noted as a general point that the approach used for probabilistic estimations of cumulative exposures is based on the RPFs. These may be derived from either NOAELs or benchmark doses (BMD). RPFs based on BMDs are a considered a scientific refinement of the hazard assessment in the basic approach. When adequate data are available to estimate a BMD, this can produce a more consistent basis for comparing relative potencies as it identifies a dose that produces a defined level of response (EFSA, 2009b). The use of the NOAEL is limited to one of the dose levels used in the study and is independent of the magnitude of any response above the NOAEL.

When determining the RPF, EFSA (2008) emphasised the need, where possible, to obtain the toxicological information from the same species under similar experimental circumstances. If different uncertainty factors have been used, EFSA (2008) suggested that such differences should be corrected prior to calculation of the RPFs.

The selection of the IC should be based on consideration of the toxicological data, which should be well defined for the common mechanism effect. Responses for common toxicity consistent with that of the CAG and the IC dose-response should be well characterised and with adequate dose-spacing between NOAEL and LOAEL (EFSA, 2009a). When selecting the IC, substances with weaker toxicity data should be avoided, as uncertainty in the data for the IC translates into uncertainty for all the individual RPFs in the CAG and hence for the combined exposure.

The assessment report should document clearly the choice of toxicological endpoint used to cumulate toxicity, the species, study design, including any differences e.g., duration of treatment, target tissue for determining common biological response, and differences in uncertainty factors used to derive reference values, or the methods to establish reference points (e.g., BMD10, BMDL10, NOAEL). The impact of all these factors on the assessment should be considered as part of the evaluation of unquantified uncertainties (section 8).

Where the reference points are accompanied by confidence intervals (e.g., for the BMD), it would be desirable in principle to incorporate them quantitatively in the probabilistic assessment, so that they are reflected in the confidence intervals for the assessment output[59]. Confidence intervals are not available for other reference points (e.g., NOAELs), and risk assessment currently is often based on the reference points rather than their confidence intervals. However, use of the BMDL has been recommended for calculating Margins of Exposure (EFSA, 2005b). Further consideration should be given to modelling the uncertainty of toxic reference points quantitatively in future. In the meantime, it should be considered subjectively (see section 8).

Where RPFs are used, an implicit assumption is that the dose response curves are parallel and that the relative potencies are applicable over the whole of the dose range. This assumption is not necessarily valid (Moser 1995) and is therefore a source of uncertainty. Whether this introduces more or less conservatism in the assessment will vary from case to case (EFSA, 2008). The use of NOAELs may represent varying response levels for different compounds, depending on dose spacing in the toxicological studies, which is another source of uncertainty that should also be discussed.

Another potentially important source of uncertainty that should be considered is the selection of substances to include in the CAG. The impact of this may be examined by repeating the assessment with and without substances whose membership of the CAG is in question.

Physiologically based modelling, i.e., PBTK and PBTD either separately or in combination, could be considered as options for refined assessment (EFSA, 2008). Such approaches are resource intensive, but may allow estimation of more relevant endpoints. In addition, these techniques are can be used to explore possible types of combined action other than dose addition.

6.2. Co-occurrence of residues

In cumulative acute dietary exposure assessments, it is necessary to take account of any correlations that may exist between the concentrations of different members of the CAG in the same food sample. Correlations could be negative, e.g., if using one member of the CAG makes it less likely that the grower will use another member of the CAG on the same crop, but could also be positive. In order to determine whether such correlations exist and, if they exist, take account of them, it is necessary to have data where the different CAG members are measured in the same samples. These issues are not relevant when assessing cumulative chronic dietary exposure, because this depends on the mean residues of each substance in each commodity and not on the particular combinations of residues present in individual samples or units of commodity.

[59] A possible approach for this is described by Van der Voet and Slob (2007).

In some residue datasets, the same substances are measured in every sample and the data are available as a complete matrix. In this situation, RPFs can be applied to combine all the substances into a single measure of cumulative potency for each sample, and the cumulative acute assessment can then be generated by applying probabilistic modelling in the same way as for a single substance assessment.

Difficulties arise when the substances analysed differ between samples so that the matrix of samples by substances is incomplete and contains a mixture of positives, non-detects and missing values.

The EU 7th Framework project ACROPOLIS (www.acropolis-eu.com) is researching various approaches to addressing incomplete residue matrices, for example, by using additional information on agricultural use of pesticide combinations. Other researchers are also developing models for this problem (e.g., Crépet and Tressou, 2011). However, these are specialised approaches that require further development and evaluation before being considered for routine use.

The Panel therefore proposes a simpler solution for **basic probabilistic assessment** of acute exposure, which captures the correlations present in the available data and replaces missing values in a way that should over-estimate the degree of positive correlation. This avoids the need to estimate or assume the level of correlation, but is conservative in the sense that it will over-estimate exposure and risk in pessimistic model runs. An unconservative alternative is provided for optimistic model runs. The proposed procedure is as follows:

- When you have a complete matrix for a particular commodity, first substitute values below LOR for each substance separately (replace them with 0 in the *optimistic model run* and the LOR in the *pessimistic model run*), then apply RPFs and model the combined 'residue' as for a single substance.

- When you have an incomplete matrix for a particular commodity:

  - For the *optimistic model runs*, substitute missing values and values below the LOR with zero for all substances, then apply RPFs and model the combined 'residue' as for a single substance.

  - For *pessimistic model runs*:

    1. Set all values below the LOR to the relevant LOR.

    2. Count the total number of missing values for each substance. Apply the methods recommended in section 4.2 for pessimistic model runs in single substance assessments to model the distribution of values of each substance for the commodity in question[60]. Use the distribution for each substance to generate as many imputed values as there are missing values for that substance. Order the generated values for each substance from high to low.

    3. Order all the samples in the data matrix for this commodity from high to low, according to their cumulative potency based on the measured data (including values set to the LOR in step 1 above, but excluding the missing values).

    4. Consider the sample with the highest cumulative potency. Fill any missing values with the highest imputed values for the relevant substances (for the same commodity).

[60] The distribution for each substance will be a mixture of values below the LOR (set to the LOR as this is a pessimistic model run) combined with a lognormal distribution of positive values. The proportion of values <LOR should be the same as that found in the measured data for the commodity in question, and the lognormal distribution should be fitted to the positive values in the measured data.

5. Proceed to the next sample. Fill any missing values with the highest remaining imputed values.

6. Repeat step 5 until you reach the end of the samples for the commodity in question. Then recalculate the cumulative potency for each sample including the measured and imputed values, and use the cumulative potencies in probabilistic acute exposure modelling as if they were sample/lot concentrations for a single substance.

7. Repeat steps 1-6 for other commodities relevant to the assessment.

A worked example of steps 1-6 is provided in Appendix 1.

These methods are necessary only for acute assessments. The approach for pessimistic model runs will apply a conservatively high correlation to impution of missing values, while retaining the observed correlation for measured values. Note that steps 1-7 relate to concentrations for samples or lots, not for individual units. In the pessimistic model runs, unit-to-unit variation should then be modelled as for single-substance assessment: this is also conservative, because it implies perfect positive correlation of unit-to-unit variability of the different substances within samples or lots, whereas in practice the correlation could be weaker or even negative[61].

The Panel recognises that setting values below the LOR equal to the LOR in pessimistic runs is likely to be very conservative and cause exceedance of reference doses in both acute and chronic cumulative assessments. However, more realistic yet still protective assumptions would vary between commodities and setting them would require in-depth analyses of use patterns, residue databases and other relevant evidence (e.g., processing effects). An example of this is provided by the US EPA's conclusion in their OP, N-methylcarbamate, and Pyrethroid cumulative assessments that sugars and syrups can be assumed not to contain any pesticide residues based in part on the extensive processing and blending that such commodities undergo, implying that values less than the LOR could reasonably be set to zero for these foods[62]. However, for commodities or foods that do sometimes contain residues, a refined assessment would be required to estimate the frequency of true zeroes, which might include refined estimation of % crop treated and/or more sophisticated modelling.

When the procedure above is used, if it would be illegal for the pesticide associated with imputed value to be used together with other substances which were measured as being present in the same sample, then a zero could be assigned in step 4 instead of the imputed value and the imputed value would be retained for the next data gap.

7. Types and format of model outputs

A wide variety of graphical and tabular outputs can be generated by probabilistic dietary exposure modelling. Common types are listed in Table 4, together with comments on the different kinds of information they provide and types of interpretation they are useful for. In all cases, care is required in the detailed design, labelling and explanation of each form of output in order to facilitate correct interpretation.

[61] It is not known how conservative the approaches described in this section will be in practice, so this methodology should be reviewed when more research and experience are available.

[62] See US EPA reports: http://www.epa.gov/opp00001/cumulative/rra-op/I_C.pdf and http://www.epa.gov/oppsrrd1/REDs/nmc_revised_cra.pdf.

Table 4: Common types of output from probabilistic dietary exposure modelling. Note that graphical and tabular summaries of input data are also useful (see section 9 on checklist of issues for reporting and peer review). Examples of some of these outputs may be found in de Boer et al. (2011).

Type of output	**Contribution to interpretation of results**
Exposure distribution:	
• Probability density function	Readily interpretable presentation of distribution shape, good for detecting presence of multiple peaks. Usual format does not show uncertainty, although this is possible.
• Cumulative density function with confidence intervals	Shows percent of population or person-days *below* any given level of exposure. Confidence intervals show quantified uncertainty. Convenient format for reading off (approximate) percentiles of the distribution.
• Exceedance function (complementary cumulative density functions) with confidence intervals	Shows percent of consumers or person-days *above* any given level of exposure. Shows quantified uncertainty. Can be useful for reading off (approximate) numerical results.
• Tables of selected distribution statistics	E.g., average and standard deviation, plus confidence intervals. Usually of less interest than percentiles.
• Exposure at specified percentiles	Selected percentiles of exposure distribution, plus confidence intervals.
• Percent of population exceeding/below a specified exposure*	Estimates of percent population or person-days above or below specified levels, e.g., ARfD or ADI, plus confidence intervals.
Contributions of different commodities to exposure	Pie charts or tables showing percentage contribution of different commodities to exposure, either aggregated over the whole population or for a specified segment (e.g., those above a specified percentile or reference value). Confidence intervals can be shown in tables but not pie charts.
Summary statistics for estimated consumption	E.g., % consumers/consumption days and mean amounts. Useful aid to understanding exposure results and to check realism of simulation.
"Drill down" of upper tail exposure estimates*	Table of information underlying individual upper tail exposures, e.g., foods contributing to the exposure, amounts consumed, residue levels, etc. Essential for assessing realism of results in upper tail and also for assessing the possibility of still higher exposures.
Sensitivity analysis	Various forms, ranging from simply presenting results of different model runs to show the impact alternative assumptions (e.g., for treatment of values below the LOR), to figures or tables showing the contribution of different inputs to variation and uncertainty in the output.

*The Panel recommends that estimates of the proportion of the population exceeding exposures of concern (see Table 5 and Figure 6) should be included in every assessment, together with drill-down information for pessimistic basic model runs and for refined assessments. Other types of output listed in this table considered as optional.

All of the formats summarised in Table 4 can contribute to communicating the results of a probabilistic assessment. Some focus on communicating different aspects of the exposure distribution while others (especially the "drill-down") focus on the essential purpose of evaluating the credibility or realism of estimated exposures (especially in the upper tail) and assist in determining what specific mitigation or other regulatory measures might be appropriate.

In presenting results, the aim should be to show what the available data and modelling can say about the incidence of different levels of exposure relative to the relevant toxic reference values, together with a clear and balanced indication of the limitations and uncertainties associated with the results. The Panel identified the following key issues and requirements:

1. Reporting results for only one or a few percentiles of the exposure distribution gives an incomplete picture and, in effect, presupposes that those percentiles are of particular interest (e.g., to risk managers). It also implies risk management judgements about the level of protection that is required (e.g., limiting attention to the 97.5th percentile would imply that exposures occurring less frequently than 1 in 40 are of no concern). Results should be reported such that the complete estimated exposure distribution can be assessed, so far as is feasible given the underlying data and modelling methodology.

2. Exposure is of interest (to risk managers and others) primarily in terms of its relation to toxic thresholds. Assessment outputs should therefore include (but need not be limited to) expression of exposure in relation to the relevant toxic reference value, e.g., as a percentage of the ARfD or ADI, or as a Margin of Exposure (MoE)[63].

3. When the results indicate potential for a proportion of exposures to exceed the relevant toxic reference value, it is important to characterise the amount by which the reference value may be exceeded, as this is critical for interpretation of the potential toxicological consequences and hence the risk management implications. Therefore, results should not be limited to estimates of the proportion of exposure exceeding the reference value, but should also estimate the proportions of exposure at different multiples of the reference value (e.g., 2x, 5x, 10x, or other multiples selected according to their potential toxicological significance). This provides an appropriate basis for toxicologists and risk managers to consider whether the requirements of Article 4.2 of Directive 1107/2009 that pesticide residues 'shall not have any harmful effects' are met.

4. It is essential to take account of uncertainty, with regard to the potential both for underestimation and overestimation of exposure. Those uncertainties that have been quantified by the assessment should be characterised by presenting confidence intervals with each estimate of exposure. In addition, any uncertainties explored by repeating the assessment with alternative assumptions (e.g., pessimistic and optimistic model runs) should be characterised by presenting results for the alternative models side by side.

5. It is equally essential to take account of uncertainties that have not been quantified. These should always be evaluated systematically (see section 8) and presented alongside the quantitative results.

6. It is also equally important (and indeed part of the assessment of unquantified uncertainties) to evaluate the credibility of the simulated exposures, especially in the upper tail of the distribution, e.g., by making use of "drill down" information. Results that are based on clearly unrealistic values or combinations of values (e.g., consumption or residues exceeding maximum feasible values, if these can be defined) may occur when

[63] Margin of Exposure is the ratio of the relevant toxicological endpoint (before application of any uncertainty factors) to an estimated exposure. Though not currently used in pesticide assessments, it is increasingly used in risk assessments relating to environmental contaminants in the diet (e.g., EFSA, 2005b).

using parametric models that extrapolate beyond the range of the observed data. In addition, it is useful to examine the input data underlying the highest exposures, and check for potential data errors (e.g., misreporting of consumption, or decimal place errors). Individual results that clearly exceed credible limits, or can be shown to derive from invalid inputs, may be omitted from the primary results presented to risk managers, provided that they are reported and their omission is justified in accompanying documents so that they are open to peer review.

7. Similarly, it is important to evaluate the potential for exposures higher than those simulated by the model, e.g., by using drill-down to examine the consumption and residues values in the upper tail of the exposure distribution and using expert judgment to assess whether higher values are likely in practice. This is essential for models based on resampling observed data as these are necessarily limited to the range of the observations, but it is also relevant for parametric models (to assess whether the tails may be too light). If these considerations indicate the potential for higher exposures, this should be clearly indicated immediately adjacent to the quantitative results.

8. A probabilistic model cannot estimate the frequency of exposures lower than the reciprocal of the size of population simulated by the model. For example, if the model simulates 100,000 person-days, it cannot generate estimated proportions lower than 1 per 100,000 (i.e., 10^{-5}, or 0.001%). In assessments where the results show credible exposures occurring at the limiting frequency of the model, the model should be rerun with a larger population. If this is not done, it is important to make clear, alongside the quantitative results, that higher exposures might occur at frequencies below the limiting frequency of the model.

To meet these requirements, the Panel recommends the use of a specific tabular and graphical format as illustrated in Table 5 and Figure 6 (adapted as appropriate for chronic or acute assessments). It is recommended that both the tabular and graphical forms are included to facilitate interpretation and understanding of the results by the reader.

It is suggested that frequencies of exposures could be expressed as percentages for assessments of high residue events, and as numbers per million for other types of assessment (approval, authorisation, MRL-setting and annual review of monitoring data, see section 3), unless risk managers prefer a different format. Percentages are suggested for high residue events because the number of people potentially exposed to a single lot of commodity will be more limited than for the wider populations considered for the other types of assessment.

The main results should always be accompanied by "drill-down" information to support a discussion of the credibility of the upper tail exposures, the validity of the underlying input data, and the potential for higher exposures to occur.

The Panel recommends that these formats (Table 5 and Figure 6, plus drill-down information and evaluation of unquantified uncertainties) be included in every **basic and refined probabilistic assessment** of dietary exposure to pesticides for regulatory purposes to provide a consistent basis for interpretation and decision-making. These may be accompanied by other formats (e.g., others from Table 4) where these are considered to provide useful additional information.

The tabular and graphical results should always be accompanied by a narrative conclusion. It should be explained that the results are estimates, and that the actual exposures are expected to lie between the optimistic and pessimistic estimates for acute exposure, and below the pessimistic estimates for chronic exposure. Where the results together with full consideration of the associated uncertainties lead to a conclusion that exposures above the toxic reference value are scientifically not credible for the pesticide use scenario under assessment, the reasoning for this should be explained. In all other cases, it should be stated that exposures

above the toxic reference value are probable or possible (according to the evidence), and the reader should be referred to the tables and figures for information on the frequency of such exposures and associated uncertainties. To aid understanding, it may be helpful to describe selected results from the tables or figures in narrative form. Interpretation of the results for decision-making is discussed in Section 10 (see below).

Table 5: Tabular format recommended by the PPR Panel for summarising results of probabilistic dietary exposure assessments. Results for optimistic and pessimistic model runs are shown side by side, and the right hand column is used to summarise any additional considerations or uncertainties that should be taken into account. The population or subpopulation to which the assessment relates should be specified in the table title. The number of individuals simulated should also be stated, together with the limiting frequency that this implies. The example is for acute assessment and shows the frequency of exposures at different multiples of the ARfD, expressed as the number of exceedances per million person-days[64]. These fictional results are supposed to have been generated by a simulation of 100,000 individuals, so the minimum frequency of exceedances that can be estimated is 10 per million. The same table format should be used for chronic assessment, but showing the frequency of exposures at different multiples of the ADI, expressed as the number of exceedances per million individuals. In assessments of high residue events (see section 3), the frequency of exceedances may be expressed as a percentage of consumption-days.

Exposure levels		**Number of person-days per million exceeding exposure level**		**Additional considerations & uncertainties[65]**
% of ARfD	**MoE***	**Optimistic model run**	**Pessimistic model run**	
1	10000	2000 (500 – 7000)	5000 (1000 – 17,000)	*Indicate overall direction and magnitude of additional uncertainties, e.g., by inserting summary text from bottom row of uncertainty table (see Section 8).*
10	1000	500 (200 – 1200)	1500 (300 – 4000)	
50	200	50 (10 – 500)	400 (100 – 1300)	
100**	100	10 (<10 – 50)	60 (20 – 300)	
200	50	<10 (<10 – 10)	10 (<10 – 40)	*Identify or omit results that are based on clearly unrealistic extremes of input distributions.*
500	20	<10 (<10 - <10)	<10 (<10 – <10)	*Use '<' to indicate results that are below the sensitivity of the model.*

<10. = lower than limiting frequency of model
parentheses = approximate 95% confidence intervals for sampling uncertainties

* Margin of Exposure assuming that the ARfD has been established with an uncertainty factor of 100 (if a different uncertainty factor has been used, this should be stated). For acute dietary exposure assessment this is the ratio of the exposure estimate to the toxicological reference point on which the ARfD is based.

** typical basis for MRL setting

[64] If the acute assessment is restricted to consumption-days only, then the results should be expressed in consumption-days rather than person-days.

[65] This information is shown in a column to allow separate comments to be made for different rows (e.g. in the upper tail). If the same comments apply to all rows, it may be more convenient to present them at the foot of the table.

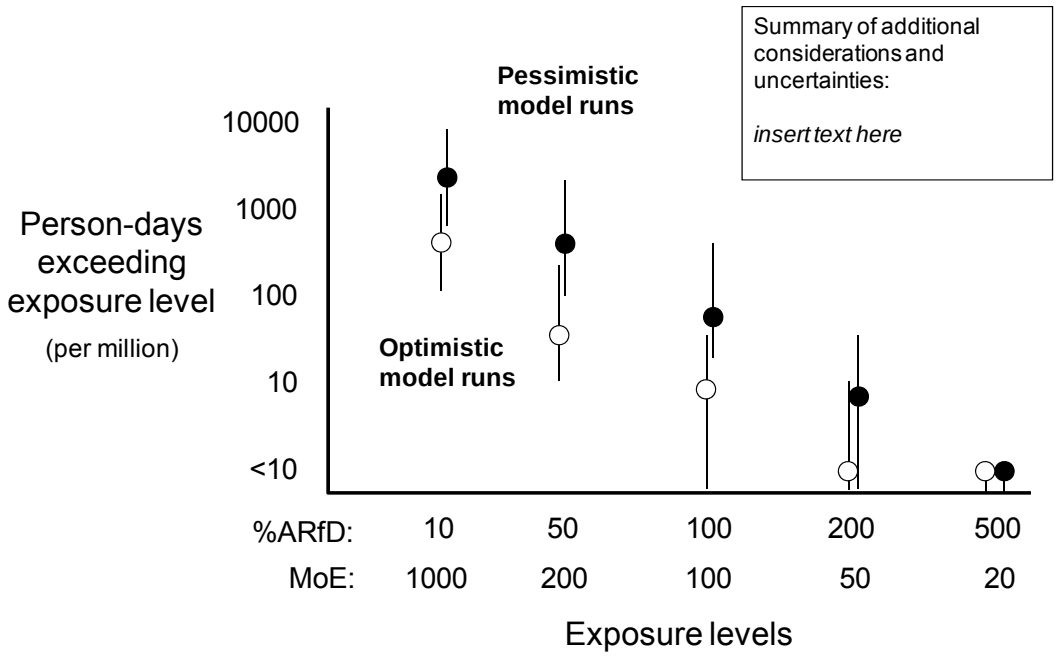


Figure 6: Graphical format recommended by the PPR Panel for summarising results of probabilistic dietary exposure assessments[66]. Results for pessimistic and optimistic model runs are shown side-by-side. The population or subpopulation to which the assessment relates should be specified in the figure title. The number of individuals simulated should also be stated, together with the limiting frequency that this implies. The example is for acute assessment and shows the frequency of exposures at different multiples of the ARfD, expressed as the number of exceedances per million person-days on a logarithmic scale[67]. These fictional results are supposed to have been generated by a simulation of 100,000 individuals, so the minimum frequency of exceedances that can be estimated is 10 per million. The same figure format should be used for chronic assessment, but showing the frequency of exposures at different multiples of the ADI, expressed as the number of exceedances per million individuals. In assessments of high residue events (see section 3), the frequency of exceedances may be expressed as a percentage of consumption-days. Vertical bars represent approximate 95% confidence intervals representing sampling uncertainty. MoE = Margin of Exposure assuming a safety factor of 100 (if a different factor is appropriate, this should be used).

8. Evaluation of uncertainties

Methods for evaluation of uncertainties in exposure assessment have been considered in detail by EFSA's Scientific Committee which recommended a tiered approach starting with simple subjective evaluation of uncertainty and progressing to deterministic or probabilistic modelling when appropriate (EFSA, 2006b).

Consistent with EFSA (2006b), the PPR Panel's recommended approaches for basic probabilistic modelling include methods for quantitative evaluation of some of key

[66] From a technical viewpoint, and to show the full distribution, a cumulative distribution (with confidence intervals) plotted on a suitable scale would be preferable, provided it is well understood by the audience.

[67] If the acute assessment is restricted to consumption-days only, then the results should be expressed in consumption-days rather than person-days.

uncertainties affecting pesticide exposure assessment. These include: (i) the use of parametric models and bootstrapping to quantify sampling uncertainty; (ii) optimistic and pessimistic model runs exploring alternative assumptions for the treatment of residues below the LOR, processing factors, empirical vs. parametric modelling of residues, and variability factors, and (iii) sensitivity analysis to examine the relative impact of different uncertainties.

However, exposure assessment is affected by many other sources of uncertainty, e.g., the relevance of the available residue monitoring data to the pesticide usage patterns at the time of the assessment, or the relevance of the standard acute and chronic time scales (one day or life time) to the time course of effects used for establishing the toxicological reference values. It is essential always to take appropriate account (in proportion to their potential importance) of all unquantified uncertainties that the assessor can identify (EFSA 2006b, EFSA 2009c). The PPR Panel therefore recommends that, **in every probabilistic assessment**, assessors should systematically examine all parts of the assessment for unquantified uncertainties and evaluate them using the tabular approach as recommended by EFSA (2006b) and illustrated in Table 6. This evaluation should initially be done for the pessimistic model run in the basic assessment, and then revised for each refined assessment (if done); a separate evaluation may optionally be done for the optimistic model run.

In order to facilitate development of an overall conclusion taking account of both the quantitative results and the unquantified uncertainties, it is recommended to define a quantitative meaning for the symbols to be used in Table 6. For example, a single – or + symbol could be defined as representing potential to cause under-estimation (-) or over-estimation (+) of the estimated exposures by a factor of around 2, two symbols might represent a factor of around 5, three symbols, a factor of around 10 and fours symbols, a factor of more than 10. Smaller or larger factors could be chosen, depending on the scale of the uncertainties to be expressed. Of course, subjective evaluation is highly approximate, and should be interpreted accordingly. Alternatively, the symbols can be interpreted more qualitatively (e.g., low, medium, high), although this is less transparent (because such terms will be interpreted differently by different people) and will make it less easy for the assessor to form an overall conclusion that combines the quantified and unquantified uncertainties.

In cases where the unquantified uncertainties appear substantial enough to alter the risk management decision, consideration may be given either to assessing them in more detail (e.g., by rerunning the model with alternative assumptions or treating them probabilistically) or to seeking additional data to reduce the uncertainty. For example, if the relevance of a 1 day timescale for acute assessment was considered doubtful, the potential importance of this could be explored by redoing the assessment with alternative timescales.

A general evaluation of the uncertainties affecting the approaches to basic probabilistic exposure assessments recommended in this guidance is presented in Appendix 2: one table for acute assessments, and the other for chronic assessments. These tables may be helpful as a starting point for users to evaluate uncertainties affecting specific assessments. If the user follows the recommended approaches for **basic probabilistic assessment** *without modification*, it may be sufficient for the assessment report to refer to the tables in Appendix 2 of this document, although the user will need to make their own assessment of the overall impact of the uncertainties (which should be considered case by case and is not provided in Appendix 2). However, as the tables in Appendix 2 take account of the wide range of situations that may be encountered in basic assessment (e.g., regarding size and quality of the data available) it may be beneficial to review the entries in the table (including the overall assessment at the foot of the table) and adjust them where appropriate to reflect the details of the assessment and data in hand.

If a basic assessment modifies the recommended approaches in any way, or in a **refined probabilistic assessment**, the tables from Appendix 2 may still be used as a starting point,

but they should be edited and re-evaluated as necessary to take account of the approaches used in the assessment.

The PPR Panel recommends that an evaluation of uncertainties using this approach should always be conducted for the *pessimistic model runs* in **basic probabilistic assessment**, and for **refined probabilistic assessments**. It is not essential to evaluate uncertainty in this way for *optimistic model runs*, since they will not be relied on for determining the acceptability of exposures (see Section 2 and Figure 1).

Table 6: Tabular approach for evaluation and expression of uncertainties affecting exposure and risk assessments (adapted from EFSA, 2006b). The +/- symbols are used by the assessor indicate whether each source of uncertainty has the potential to increase (+) or decrease (-) the assessment outcome. The number of symbols represents the assessor's subjective evaluation of the magnitude of the effect (e.g., +, ++ and +++ might indicate uncertainties that could cause over-estimation by x2, x5 and x10 respectively). If the effect is uncertain, or could vary over a range, lower and upper evaluations are given (e.g., - / ++ or + / ++). Finally, the combined impact of all the uncertainties is evaluated subjectively. More detail on the rationale for these evaluations (especially for the more important uncertainties and the overall uncertainty) should be provided as separate text accompanying the table. See Appendix 2 for completed Tables for uncertainties affecting acute and chronic exposure assessments conducted according to the recommendations in this guidance.

Source of uncertainty	**Magnitude and direction of influence on estimated exposures**
Concise description of source of uncertainty (e.g., under-reporting of consumption of some commodities in dietary surveys)	Symbols to show evaluation of influence (e.g.: +/++)
Insert one row for each source of uncertainty affecting the assessment	
Overall evaluation of uncertainty affecting the assessment outcome Add narrative text here, describing the assessor's subjective evaluation of the overall degree of uncertainty affecting the assessment outcome, taking account of all the uncertainties identified above.	Evaluation of overall uncertainty (e.g., - - - /+)

9. Checklist of key issues for report-writing and peer reviewers

This section summarises key issues that assessors should address when producing reports on probabilistic dietary exposure assessments, and which specialists evaluating or peer-reviewing those reports should check for.

Purpose and scope of assessment

1. The purpose of the assessment, the scientific and/or regulatory questions it addresses, its focus (pesticide uses and commodities) and scope (acute or chronic exposure, population of interest).

Input data

2. Descriptions of all the data used as inputs for the assessment (including consumption and residue data, food conversion factors, unit weights, processing factors, etc.), justification

of their relevance to the assessment, and references to detailed information on how the data were collected and where they can be found. The description should be detailed enough to enable an independent reviewer to replicate the assessment precisely.

3. Summary statistics (mean, standard deviation, range, sample size) for each set of input data, presented separately for each commodity. In addition, histograms should be provided for commodities which contribute significantly to the overall exposure, to allow examination of distribution shapes.

4. A description of how composite food as eaten is converted to individual ingredients.

5. Table listing the RACs relevant for the assessment, showing which are modelled using residue data from monitoring programmes, which are modelled using data from supervised field trials/feeding studies, and which are modelled by extrapolation from other RACs.

Distributional assumptions

6. List of all parametric distributions used in the assessment, identifying the data on which they are based and how they were estimated. Graphical and statistical assessment of appropriateness and goodness of fit for each parametric distribution used.

Model structure

7. Description of the model structure, including distributions and equations, sufficient for it to be reproduced by others, or reference to a published source where this information can be found.

8. Justification of appropriateness of the model for the purposes of the assessment, or reference to a published source where its suitability for this purpose is documented (e.g., validation studies, see section 11).

9. Size of population modelled and demonstration that this is sufficient to produce stable outputs including at tail percentiles relevant for decision-making.

10. Number of bootstrap iterations performed, and demonstration that this is sufficient to produce stable confidence intervals at tail percentiles relevant for decision-making.

Software

11. Identity (including version numbers) and description of software used, or reference to published sources where this information can be found.

12. A list of all software settings used in the assessment, sufficient for it to be reproduced.

Refined probabilistic assessments

13. Full description of any refinements of the assessment beyond the basic probabilistic approaches specified in this guidance, sufficient for them to be exactly reproduced by others, together with full scientific and statistical justification of their appropriateness.

Outputs

14. Outputs in the form of Table 5, Figure 6, drill down information (for evaluating the realism of estimated exposures in the region relevant for decision-making), Table 6 (evaluation of unquantified uncertainties) and a narrative summary conclusion should be provided for every assessment.

15. Tables and graphs showing contribution of different food items to the total exposure and to exposures above the relevant toxicological reference value (ARfD or ADI).

16. Comparison of means, medians and quartiles for measured and simulated consumption and residue levels (as a check on the quality of the simulation).

17. Optionally, other outputs such as those listed in Table 4, if these contribute to understanding of the assessment and its results.

Uncertainties

18. List of uncertainties quantified by bootstrapping, parametric modelling and sensitivity analysis (including rerunning model with alternative assumptions).

19. Results and interpretation of any sensitivity analyses.

20. An evaluation of unquantified uncertainties in the form of Table 6 should be provided for every assessment, including assessment of their potential overall impact on the estimated exposures.

Justification of exceptions and deviations

21. Justification for any exceptions and deviations from the recommendations in this guidance, and evaluation of their impact on the assessment outcome.

10. Interpretation of results and options for risk managers

As stated earlier, the aim in presenting the results should be to show what the available data and modelling can say about the incidence of different levels of dietary exposure relative to the relevant toxic reference values, together with a clear and balanced indication of the limitations and uncertainties associated with the results. The aim of the Panel's recommendations for both the conduct of the assessment and presentation of the results is to provide a sound basis for consideration by risk managers.

As probabilistic assessments provide new types of information, the following comments may be helpful support to decision-making:

- Table 5 and Figure 6 are designed to provide the key information relevant for decision-making: estimates of the frequency of exposures exceeding toxic reference values, and the quantified and unquantified uncertainty associated with these estimates.

- In generating these outputs, the assessor will examine the realism of the results and highlight and discuss any that are clearly beyond what is scientifically credible.

- When communicating results, it should be clearly explained that the results are estimates, and that the actual exposures are expected to lie between the optimistic and pessimistic estimates for acute exposure, and below the pessimistic estimates for chronic exposure.

- The degree of uncertainty associated with the assessment is indicated by the difference between the results for optimistic and pessimistic model runs, the confidence intervals for quantified uncertainties, and the subjective evaluation of unquantified uncertainties (evaluated in Table 6 and summarised in the right hand column of Table 5). These should be considered together when forming overall conclusions. If the unquantified uncertainties are minor or negative, then it is very likely that the true exposure

distribution lies below the upper confidence interval for the pessimistic run and (in acute assessments) probably above the lower confidence interval for the optimistic run. If the unquantified uncertainties are large and positive, this indicates the potential for exposures above the upper confidence interval for the pessimistic run.

- If the estimated exposures extend above the toxic reference values, toxicologists may assist in considering the toxicological implications of those exposures: what types of effects may occur, and at what levels of exposure are they to be expected?

- Possible considerations for risk managers include: the estimated degree and frequency of exceedances of toxic reference values (if any), the nature and frequency of adverse effects expected (if any), and the degree of uncertainty in these estimates. The question for risk managers to consider is: does the assessment provide adequate certainty that the frequency and severity of effects will not exceed a level of concern? Note that the acceptable frequency and degree of certainty required may vary, depending on the severity of the effects in question[68].

- In cases where risk managers wish to reduce exposures, it may be helpful to examine assessment outputs showing which pesticide/commodity combinations contribute most to the overall exposures, as this may help in identifying options for exposure reduction.

- In cases where risk managers wish to reduce uncertainty, it may be helpful to examine outputs showing which pesticide/commodity combinations contribute most to overall uncertainty, as this may help in identifying priorities for additional data collection.

Good communication between the authors of the probabilistic assessment and the risk managers is important to facilitate interpretation of the results, especially considering that these approaches are relatively new and can be complex. In cases where risk managers wish to consider requesting refined probabilistic assessment and/or new data collection, they may benefit from advice on the feasibility and potential usefulness of these options for the case in hand.

11. Validation

Before accepting the results of probabilistic modelling, it is essential to consider how well the modelled dietary exposures predict the actual exposures for the scenario assessed.

In a previous Opinion, the PPR Panel compared acute exposures estimated with the Monte Carlo Risk Assessment (MCRA) model (version 5.1) with measured one-day exposures from duplicate diet studies conducted by Boon et al. (2003a) and López et al. (2003). Although these duplicate diet studies did not provide complete distributions of exposures, the results showed that the measured exposure at the 99th percentile was a factor of 10 or more below the modelled exposure based on residue data from monitoring, which in turn was lower than the modelled exposure using residue data from supervised field trials/feeding studies.

Based on these results the Panel concluded that there is some support for thinking that probabilistic modelling using concentrations from monitoring programmes may tend to over-estimate true exposures (EFSA, 2007a). However, the Panel noted that the comparisons were done for infants, and there are factors that might make the difference between measured and modelled exposure smaller for other age groups. In addition, the comparisons were limited to

[68] Determining the level of concern is a matter for risk managers. Article 4.2 of Directive 1107/2009 requires that residues of plant protection products 'shall not have any harmful effects on human health'.

6 pesticides in one country and it is uncertain how representative they are of other pesticides and other countries.

Respondents can only be asked to collect duplicate diets for a couple of days and consequently for practical reasons duplicate diets are not suitable to validate chronic exposure assessment. Some chronic exposure models have been compared with biomarkers rather than direct measurement of consumption or exposure. Biomarkers have been used as a measure of exposure to pesticides, but include exposure also via other routes than food only, and therefore these studies may not be suitable for validation.

Slob et al. (2010) have used a computer simulation of actual exposures to evaluate the performance of exposure models. The simulation model generates distributions of exposure to a chemical via two foods. The simulated exposure distribution was compared with the exposure distributions estimated using the BBN or the ISUF model. Given the practical difficulties of measuring actual exposures empirically, simulation models of this type may provide useful tools for evaluating and improving the currently available dietary exposure models.

12. Software quality requirements

The PPR Panel does not endorse any particular software for dietary exposure modelling, but offers the following criteria for consideration by users when deciding which software to use.

1. The software should be able to carry out exposure assessments following the approaches recommended in this document, including the alternative assumptions and modelling methods used to explore key uncertainties.

2. Software should generate (or enable the user to generate) graphical and tabular outputs of the types discussed in section 6, including in particular outputs in the format of Tables 3 and 4 and Figure 3 and drill-down information.

3. The software should include an openly available reference manual describing the statistical models used in order to be transparent and in order to enable third parties to reproduce the results.

4. The software should preferably permit both short-term and long-term exposure assessment.

5. The software should include methods for quantifying uncertainty.

6. Simulations should be performed within reasonable time.

7. Any data which are incorporated within the software should be documented and justified.

8. The software should be designed to avoid any double-counting in its treatment of food conversion factors and processing factors (see section 4.3).

9. Software should preferably be freely available or at least available without substantial monetary or non-monetary barriers.

10. The possibility for pesticide industry, regulatory authorities and other stakeholders to use the same model and data without the user having to repeat the data input and model set-up, e.g., by running on the internet or by provision of downloadable models with datafiles that can be exchanged between users.

CONCLUSIONS

Future developments

The approaches recommended in this document take account of the current state of the art including practical methods for addressing uncertainties affecting the data and models used in probabilistic dietary exposure assessment. They include sensitivity analysis to help identify key areas of uncertainty, so that these can be considered in refined assessment where appropriate. Expert judgment is often required, especially in refined assessments.

The approaches recommended in this document can in principle be applied immediately by users with relevant data and modelling expertise. It is anticipated that those aspects that are not currently implemented in ready-to-use software are likely to be added in the near future. However, further work is required to make the methods available to the end-user in practical form.

Users would benefit from easier arrangements for access to the necessary data, and especially so if relevant data from consumption surveys and residues would be made available in a central database in consistent form. Users would also benefit from organised provision of training in the principles and conduct of probabilistic assessment.

Further work could be undertaken to further evaluate and refine the approaches for probabilistic exposure assessment, and for addressing the key uncertainties identified in this document. Such work could provide approaches for use in refined assessments, and/or for possible future revisions of this guidance. Areas that might benefit from further work include:

1. Carry out simulation studies to examine the performance of alternative approaches for handling variation of residues between and within lots or samples, including the factor to be used to account for unit-to-unit variability, and take account of the results when finalising the choice of methodology to use in basic probabilistic assessments of acute exposure.

2. Development or adaptation of software to make the proposed approaches for basic probabilistic assessments available to users in a friendly and practical form, with appropriate user guidance.

3. Make available, in organised, quality-assured and accessible form, the data needed for probabilistic assessments, including consumption data, residues from monitoring and field trials, processing factors, food conversion factors and the registration status of substances and uses.

4. Additional case studies to demonstrate and evaluate the approaches recommended in this document.

5. Develop more specific guidance and methods for refined assessment, including:

    a. Improved approaches for modelling variation of residues between lots or samples, especially in the upper tails.

    b. Improved understanding of the nature and importance of variation of unit residues within lots or samples, including the mixing of treated and untreated commodity, and modelling strategies to take account of this in acute exposure assessments.

    c. Options for taking account of uncertainty arising from parameters for which only very few data are available, including processing factors.

d. Further improvement and evaluation of methods for modelling habitual consumption or exposure for chronic exposure assessments, including methods suitable for assessments involving multiple food types.

e. Modelling of repeated acute exposures and/or exposures over time periods intermediate between acute and chronic, if required by risk managers.

f. Further investigation of the derivation and use of conversion factors for residue definition in risk assessment.

6. Consultation with risk managers when more case studies are available, to explore the feasibility of developing thresholds for use in deciding when to conduct probabilistic assessments and how to interpret the results, and to consider how these thresholds should take account of factors such as the severity of effects and the level of uncertainty.

7. Validation studies based on either duplicate diet studies, biomarker studies, or simulation studies.

REFERENCES

Allcroft DJ, Glasbey CA, Paulo MJ, 2007. A latent Gaussian model for multivariate consumption data. Food Quality and Preference, 18: 508-516.

Boon PE, Van der Voet H and Van Klaveren JD, 2003a. Validation of a probabilistic model of dietary exposure to selected pesticides in Dutch infants. Food Additives and Contaminants 20 Suppl 1: 36-49.

Boon PE, Van der Voet H and Van Klaveren JD, 2003b. Dietary exposure to pesticides - relevant variables and probabilistic models. RIKILT report 2003-008. Wageningen.

Boon PE and Van Klaveren JD, 2003. Guidelines regarding probabilistic exposure assessment in the safety evaluation of pesticides in the EU Market. Study on Probabilistic Assessment Consumer Training. Rikilt Institute of Food Safety, B1-3330/SANCO/2002584.

Bosgra S, Van der Voet H, Boon PE, and Slob W, 2009. An integrated probabilistic framework for cumulative risk assessment of common mechanism chemicals in food: An example with organophosphorus pesticides. Regulatory Toxicology and Pharmacology (54) 124-133.

Breuninger MI, Schachtel GA, Lührmann PM, Hartmann B, Neuhäuser-Berthold M, 2003. Complex correlation structure in consumption rates of major food groups: implications for the assessment of radiation exposure. Public Health Nutrition: 6, 87–94.

Codex Committee on Pesticide Residues, 2002. The Probabilistic approach to acute dietary exposure analysis and its applicability at the international level (Annex). Thirty-fourth Session, The Hague, 13-18 May 2002.

Crépet A and Tressou J, 2011. Bayesian nonparametric model for clustering individual co-exposure to pesticides found in the French diet. Bayesian Analysis, 6:127-146.

Cullen AC and Frey HC, 1999. The use of probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs. New York, NY, Plenum Press.

de Boer, WJ, Van der Voet H, 2011. MCRA 7.1 Reference Manual. Report Dec 2011. Biometris and RIVM, available at https://mcra.rivm.nl

de Boer WJ, Van der Voet H, Bokkers BGH, Bakker MI, Boon PE, 2009. A comparison of two models for estimating usual intake addressing zero consumptions and non-normality. Food Additives and Contaminants. 26: 1433-1449.

EC (European Commission, Health & Consumer Protection Directorate-General), 2008. Guidance Document Guidelines on comparability, extrapolation, group tolerances and data requirements for setting MRLs, SANCO 7525/VI/95 - rev.8, 1 February 2008.

Efron, B and Tibshirani, R, 1993. An Introduction to the Bootstrap. Boca Raton, FL: Chapman & Hall/CRC.

EFSA (European Food Safety Authority), 2005a. Opinion of the Scientific Panel on Plant health, Plant protection products and their Residues on a request from Commission related to the appropriate variability factor(s) to be used for acute dietary exposure assessment of pesticides in fruit and vegetables. The EFSA Journal, 177, 1-61.

EFSA (European Food Safety Authority), 2005b. Opinion of the Scientific Committee on a request from EFSA related to A Harmonised Approach for Risk Assessment of Substances Which are both Genotoxic and Carcinogenic. The EFSA Journal, 282, 1-31.

EFSA (European Food Safety Authority), 2006a. Opinion of the Scientific Panel on Plant health, Plant protection products and their Residues on a request from the Commission related to the revision of Annexes II and III to Council Directive 91/414/EEC concerning

the placing of plant protection products on the market – Residues. The EFSA Journal, 364, 1-27.

EFSA (European Food Safety Authority), 2006b. Guidance of the Scientific Committee on a request from EFSA related to Uncertainties in Dietary Exposure Assessment. The EFSA Journal, 438, 1-54.

EFSA (European Food Safety Authority), 2007a. Opinion of the Scientific Panel on Plant protection products and their Residues on a request from the Commission on acute dietary intake assessment of pesticide residues in fruit and vegetables. The EFSA Journal, 538, 1-88.

EFSA (European Food Safety Authority), 2007b. Reasoned opinion on the potential acute and chronic risk to consumers' health arising from temporary MRLs.

EFSA (European Food Safety Authority), 2008. Opinion of the Scientific Panel on Plant Protection products and their Residues to evaluate the suitability of existing methodologies and, if appropriate, the identification of new approaches to assess cumulative and synergistic risks from pesticides to human health with a view to set MRLs for those pesticides in the frame of Regulation (EC) 396/2005. The EFSA Journal, 704, 1-84.

EFSA (European Food Safety Authority), 2009a. Scientific Opinion on risk assessment for a selected group of pesticides from the triazole group to test possible methodologies to assess cumulative effects from exposure throughout food from these pesticides on human health on request of EFSA. EFSA Journal, 7 (9): 1167 (104pp.).

EFSA (European Food Safety Authority), 2009b. Use of the benchmark dose approach in risk assessment. Guidance of the Scientific Committee. The EFSA Journal, 1150, 1-72.

EFSA (European Food Safety Authority), 2009c. Guidance of the Scientific Committee on Transparency in the Scientific Aspects of Risk Assessments carried out by EFSA. Part 2: General Principles. The EFSA Journal (2009), 1051, 1-22.

EFSA (European Food Safety Authority), 2010a. 2008 Annual Report on Pesticide Residues according to Article 32 of Regulation (EC) No 396/2005. EFSA Journal 2010; 8(7):1646 [442 pp.].

EFSA (European Food Safety Authority), 2010b. Management of left-censored data in dietary exposure assessment of chemical substances. EFSA Journal 2010; 8(3):1557.

EFSA Panel on Plant Protection Products and Their Residues (PPR), 2012. Scientific Opinion on evaluation of the toxicological relevance of pesticide metabolites for dietary risk assessment. EFSA Journal 2012; 10(7):2799

European Commission (1997). Guidelines for the generation of data concerning residues as provided in Annex II part A, section 6 and Annex III, part A, section 8 of Directive 91/414/EEC concerning the placing of plant protection products on the market: Appendix E - Processing studies, (Doc. 7035/VI/95 rev. 5), 22 July 1997. http://ec.europa.eu/food/plant/protection/pesticides/docs/app-e.pdf

FOAM 2009 (second edition) Food and Agriculture Organization of the United Nations, Submission and evaluation of pesticide residue data for the estimation of maxium residue levels in food and feed.

Goedhart PW, van der Voet H, Knüppel S, Dekkers ALM, Dodd, KW, Boeing H, van Klaveren JD (2012). A comparison by simulation of different methods to estimate the usual intake distribution for episodically consumed foods. Report. Biometris, Wageningen University and Research centre.

Hill A R, 2000. Residue variability and sampling – practical problems and consequences for residues monitoring. Food Additives and Contaminants, 17, 539-546.

Hill AR, Reynolds SL, 2002. Unit-to-unit variability of pesticide residues in fruit and vegetables. Food Addit Contam. 19:733-47.

Horwitz W. 1982. Evaluation of analytical methods for regulation of food and drugs. Analytical Chemistry 54:67A–76A.

IPCS/WHO (International Program of Chemical Safety/World Health Organization), 2008. Uncertainty and data quality in exposure assessment. Harmonisation Project Document No. 6. WHO, Geneva.

ISEA (International Society for Exposure Analysis), 2005. Zartarian V, Bahadorf T and McKone T, 2005. Adoption of an Official ISEA Glossary. J. Expos. Anal. & Environ. Epid. 15 (1-5).

IUPAC (International Union of Pure and Applied Chemistry), 2006. Stephenson, Gerald R, Ferris I G, Holland P T and Nordgerg M, 2006. Glossary of terms relating to pesticides, (IUPAC Recommendations, 2006), Pure Appl. Chem., Vol. 78, No. 11, pp. 2075–2154.

JMPR, 2003. Pesticide residues in food 2003. Joint FAO/WHO Meeting on Pesticide Residues, http://www.fao.org/ag/AGP/AGPP/Pesticid/JMPR/Download/2003.pdf

Kennedy MC, 2010. Bayesian modelling of long-term dietary intakes from multiple sources. Food and Chemical Toxicology. 48, 250–263.

Kennedy MC and Hart A, 2009. Bayesian modelling of measurement errors and pesticide concentrations in dietary risk assessments. Risk Analysis, 29, 1427 – 1442.

Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, Subar AF, Tooze JA, Carroll RJ, Freedman LS (2009). Modeling data with excess zeros and measurement error: Application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics., 65: 1003-1010.

López A, Rueda C, Armentia A, Rodriguez M, Cuervo and Ociuo L A, 2003. Validation and sensitivity analysis of a probabilistic model for dietary intake assessment to pesticide residues with a Basque Country duplicate diet study. Food Additives and Contaminants 20 Suppl 1: 87-101.

Moser VC, 1995. Comparisons of the acute effects of cholinesterase inhibitors using a neurobehavioral screening battery in rats. Neurotoxicology and Teratology, 17, 617-625.

Nusser SM, Carriquiry AL, Dodd KW, Fuller WA (1996). A semiparametric transformation approach to estimating usual daily intake distributions. J Am Stat Ass. 91: 1440-1449.

Nusser SM, Fuller WA, Guenther PM, 1997. Estimating usual dietary intake distributions: adjusting for measurement error and nonnormality in 24-hour food intake data. In: Lyberg L, Biemer P, Collins M, DeLeeuw E, Dippo C, Schwartz N, Trewin D (editors). Survey Measurement and Process Quality. New York (NY), Wiley. p. 689-709.

OECD 2008, Magnitude of the Pesticide Residues in Processed Commodities, OECD Guidelines For The Testing Of Chemicals 508. http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-5-other-test-guidelines_20745796

OECD (Organization for Economic Co-operation and Development), 2009. Guidance document on definition of residue (as revised in 2009). OECD Series on Pesticides, No 31: Series on testing and assessment No 64.

OECD (Organization for Economic Co-operation and Development), 2011. OECD MRL Calculator: Statistical White Paper. Series on Pesticides, No. 57. ENV/JM/MONO(2011)3.

O'Hagan A, Buck, CE, Daneshkhah A, Eiser JE, Garthwaite PH, Jenkinson DJ, Oakley JE and Rakow T, 2006. *Uncertain Judgements: Eliciting Expert Probabilities*. John Wiley and Sons.

Ott WR, 1995. Environmental Statistics and Data Analysis. Lewis Publishers.

Paulo MJ, Van der Voet H, Wood JC, Marion GR, Van Klaveren JD, 2006. Analysis of multivariate extreme intakes of food chemicals. Food and Chemical Toxicology, 44: 994-1005

Pieters MN, Ossendorp BC, Bakker M I and Slob W, 2005. Probabilistic modelling of dietary intake of substances. The risk management question governs the method. RIVM report 320011001/2005.

Renwick AG, Barlow SM, Hertz-Picciotto I, Boobis AR, Dybing E, Edler L, Eisenbrand G, Grieg JB, Kleiner J, Lambe J, Muller DJG, Smith MR, Tritscher A, Tuijtelaars S, Van den Brandt PA, Walker R, Kroes R, Risk characterization of chemicals in food and diet. Food Chem. Toxicol. 2003;41:1211-1271

SANCO, 2011a. Method validation and quality control procedures for pesticide residues analysis in food and feed. Document No. SANCO/12495/2011.

SANCO, 2011b. Guidance Document SANCO 7525/VI/95, Appendix D. Guidelines on comparability, extrapolation, group tolerances and data requirements for setting MRLs. http://ec.europa.eu/food/plant/protection/pesticides/publications_en.htm

Slob W, 1993. Modeling long-term exposure of the whole population to chemicals in food. Risk Anal. 13: 525-530.

Slob W, 2006. Probabilistic dietary exposure assessment taking into account variability in both amount and frequency of consumption. Food Chem Toxicol. 44: 933-951. Corrigendum on in 44:1952.

Slob W, de Boer WJ and Van der Voet H, 2010. Can current dietary exposure models handle aggregated intake from different foods? A simulation study for the case of two foods. Food Chem. Toxicology,48, 178-186.

Tooze JA, Midthune D, Dodd KW, Krebs-Smith SM, Subar AF, Carroll RJ and Kipnis V, 2006. A New Statistical Method for Estimating the Distribution of Usual Intake of Episodically Consumed Foods, Journal of the American Dietetics Association, 106, 1575-87.

U.S. EPA (United States Environmental Protection Agency), 1996, online. Residue Chemistry Test Guidelines: OPPTS 860.1520 Processed Food/Feed. EPA 712–C–96–184. August. Available at http://www.deepwater.org/tech-users/860-1520.pdf.

U.S. EPA (United States Environmental Protection Agency), 1997a, online. Risk Assessment Forum. Guiding Principles for Monte Carlo Analysis. EPA/630/R-97/001. Available at http://www.epa.gov/raf/publications/pdfs/montecar.pdf

US EPA, 1997b. Terms of Environment: Glossary, Abbreviations and Acronyms. [online] Washington DC.

Van der Voet H, Boon PE and van Klaveren JD, 2003. Validation of Monte Carlo models for estimating pesticide intake of Dutch infants. Report 2003.002, RIKILT. Available at http://www.rikilt.wur.nl/NL/publicaties/Rapporten and http://montecarlo.tchpc.tcd.ie/reports/pdf/RIKILT%20WP4%20report.pdf. Appendix 3. Unit variability in probabilistic models for pesticide exposure assessment

Van der Voet H, Van der Heijden GWAM, Bos PMJ, Bosgra S, Boon PE, Muri SD and Brüschweiler BJ, 2009. A model for probabilistic health impact assessment of exposure to food chemicals. Food and Chemical Toxicology, 47, 2926-2940.

Van der Voet H, and Slob W, 2007. Integration of probabilistic exposure assessment and probabilistic hazard characterization. Risk Anal. 27: 351-371.

van Klaveren JD and Boon P, 2009. Probabilistic risk assessment of dietary exposure to single and multiple pesticide residues or contaminants: Summary of the work performed within SAFE FOOD project. Food and Chemical Toxicology, 47, 2879-2882

van Klaveren JD, Van Donkersgoed G, Van der Voet H, Stephenson C and Boon PE, 2010. Scientific / Technical Report submitted to EFSA. Cumulative Exposure Assessment of Triazole Pesticides. RIKILT – Institute of Food Safety, Wageningen University and Research Centre.

Vose D, 2008. Risk Analysis: A Quantitative Guide. Third edition. John Wiley and Sons.

WHO (World Health Organization), 1997a. Guidelines for predicting dietary intake of pesticide residues (revised).

WHO (World Health Organisation), 1997b. Food consumption and exposure assessment of chemicals

Zhang S, Krebs-Smith S, Midthune D, P_erez A, Buckman DW, Kipnis V, Freedman LS, Dodd KW, Carroll RJ (2011). Fitting a Bivariate Measurement Error Model for Episodically Consumed Dietary Components. International Journal of Biostatistics, 7(1), Article 1.

APPENDIX 1. WORKED EXAMPLE OF APPROACH FOR REPLACEMENT OF MISSING RESIDUE VALUES FOR CUMULATIVE EXPOSURE ASSESSMENT

This worked example shows how to replace missing residue values for use in the pessimistic cumulative exposure, using the approach described in Section 6.2. The example shows a case where the RPF method is used. If this is not possible e.g., because of lack of RPF's it will also be possible to use the HI or aHI (EFSA (2008). In these cases the RPF in the tables are substituted with ARfDs and the CR is calculated by dividing the concentrations with the ARfD.

The first table below shows hypothetical results of analysis of four different samples (S1, S2, S3 and S4) of a commodity for four different pesticides (P1, P2, P3 and P4). Each sample was analysed for only some of the pesticides. The results are given in mg/kg.

	S1	S2	S3	S4	RPF
P1	0.1	0.5	0.35	0.9	1
P2	n.a	n.a	0.6	0.2	3
P3	<LOR	0.8	0.7	n.a	0.5
P4	0.4	n.a	n.a	n.a	2

n.a = not analysed
<LOR: Below reporting limit

Step 1. Set all values below the LOR to the LOR.
There is one value < LOR. The LOR is 0.05 mg/kg for P3, so this value is assigned for P3 in S1.

	S1	Sample 2	Sample 3	Sample 4	RPF
P1	0.1	0.5	0.35	0.9	1
P2	n.a	n.a	0.6	0.2	3
P3	**0.05**	0.8	0.7	n.a	0.5
P4	0.4	n.a	n.a	n.a	2

Step 2.
a) Count the total number of missing values for each substance.
b) Apply the methods recommended in section 4.2 for pessimistic model runs in single substance assessments to model the distribution of values of each substance for the commodity in question.
c) Use the distribution for each substance to generate as many imputed values as there are missing values for that substance.
d) Order the generated values for each substance from high to low.

The imputed values generated from the distributions are shown in the table below (the working for this is not presented here). Since P2 has two missing values, two values are generated for this pesticide. The procedure is repeated for P3 (1 missing value) and P4 (3 missing values).

	S1	S2	S3	S4	RPF	Generated values
P1	0.1	0.5	0.35	0.9	1	***None***
P2	n.a	n.a	0.6	0.2	3	***0.4, 0.2***
P3	*0.05*	0.8	0.7	n.a	0.5	***0.45***
P4	0.4	n.a	n.a	n.a	2	***0.7, 0.5, 0.3***

Step 3. Order all the samples in the data matrix for this commodity from high to low, according to their RP based on the measured data (including values set to the LOR in step 1 above).

Cumulative potency (CR) = RPF1xC1 + RPF2xC2 + RPF3xC3 etc.

	S1	S2	S3	S4	RPF	Generated values
P1	0.1	0.5	0.35	0.9	1	*None*
P2	n.a	n.a	0.6	0.2	3	*0.4, 0.2*
P3	*0.05*	0.8	0.7	n.a	0.5	*0.45*
P4	0.4	n.a	n.a	n.a	2	*0.7, 0.5, 0.3*
CR	***0.925***	***0.9***	***2.5***	***1.5***		
Order of CR	***3***	***4***	***1***	***2***		

RP for Sample1 = 1x0.1 + *0.5x0.05* + 2x0.4 = 0.1+0.025+0.8=0.925
RP for Sample2 = 1x0.5 + 0.5x0.8 = 0.9
RP for Sample3 = 1x0.35 + 3x0.6 + 0.5x0.7= 2.5
RP for Sample4 = 1x0.9 + 3x0.2 = 1.5

Step 4. Consider the sample with the highest RP. Fill any missing values with the highest imputed values for the relevant substances.

S3 has the highest RP and one value is missing, for P4. Replace this with the highest value measured from distribution for P4, i.e., 0.7 mg/kg.

	S1	S2	S3	S4	RPF	Generated values
P1	0.1	0.5	0.35	0.9	1	*None*
P2	n.a	n.a	0.6	0.2	3	*0.4 , 0.2*
P3	*0.05*	0.8	0.7	n.a	0.5	*0.45*
P4	0.4	n.a	***0.7***	n.a	2	*(0.7), 0.5, 0.3*
CR	*0.925*	*0.9*	*2.5*	*1.5*		
Order of CR	*3*	*4*	*1*	*2*		

Step 5. Proceed to the next sample. Fill any missing values with the highest remaining imputed values.
S4 has the second highest RP with missing values for P3 and P4.

	S1	S2	S3	S4	RPF	Generated values
P1	0.1	0.5	0.35	0.9	1	*None*
P2	n.a	n.a	0.6	0.2	3	*0.4 , 0.2*
P3	*0.05*	0.8	0.7	***0.45***	0.5	*(0.45)*
P4	0.4	n.a	*0.7*	***0.5***	2	*(0.7), (0.5), 0.3*
CR	*0.925*	*0.9*	*2.5*	*1.5*		
Order of CR	*3*	*4*	*1*	*2*		

Step 6.
a) Repeat step 5 until you reach the end of the samples for the commodity in question.
b) Then recalculate the RP for each sample including the measured and imputed values
c) Use the CPs in probabilistic acute exposure modelling as if they were sample/lot concentrations for a single substance.

	S1	S2	S3	S4	RPF	Generated values
P1	0.1	0.5	0.35	0.9	1	*None*
P2	***0.4***	***0.2***	0.6	0.2	3	*0.4, 0.2 – all used*
P3	*0.05*	0.8	0.7	*0.45*	0.5	*0.45 – used*
P4	0.4	***0.3***	*0.7*	*0.5*	2	*0.7, 0.5, 0.3 – all used*
CR	*0.925*	*0.9*	*1.6*	*1.5*		
Order of CR	*3*	*4*	*1*	*2*		
CR recalculated	***2.1***	***2.1***	***3.9***	***2.7***		

RP for Sample1 = 1x0.1+*3x0.4*+*0.5x0.05*+2x0.4 = 0.925 + 1.2 = 2.125
RP for Sample2 = 1x0.1+*3x0.2*+0.5x0.8+*2x0.3* = 0.9 + 0.6 + 0.6 = 2.1
RP for Sample3 = 1x0.35+ 3x0.6+0.5x0.7+*2x0.7*= 3.9
RP for Sample4 = 1x0.9+3x0.2+*0.5x0.45+2x0.5* = 1.5 + 0.225 + 1 = 2.725

Step 7. Repeat steps 1-6 for other commodities relevant to the assessment.

APPENDIX 2. EVALUATION OF UNCERTAINTIES FOR THE APPROACHES RECOMMENDED IN THIS GUIDANCE

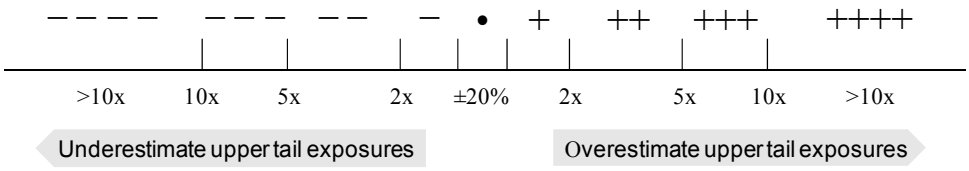
This appendix uses the uncertainty table format to assess uncertainties affecting pessimistic model runs in basic probabilistic assessment, when conducted according the recommendations in this guidance. This is important because it evaluates the basis for risk managers to be confident that decisions based on the pessimistic model would indeed be protective. Table 7 provides the Panel's evaluation of the approaches for acute exposure, while Table 8 evaluates the approaches for chronic exposure. It was decided to evaluate the *impact of uncertainties on the highest simulated exposures*, since these are likely to be a key focus for consideration in decision making (see section 10).

The Panel defined a scale for the evaluation of the uncertainties, as shown in the legend to Table 7. Pluses indicate that an uncertainty may be causing the probabilistic model outputs to be overestimates, i.e., both the median estimates and confidence intervals may be too high. Minuses mean the probabilistic model outputs may be underestimates, i.e., both the median estimates and confidence intervals may be too low. Where both pluses and minuses are shown, this means the median estimates could be too high or too low (depending on the balance of plus and minus symbols) and the confidence intervals could be wider in both directions. Thus the symbols can be interpreted as indicating how the extra uncertainties might alter the median estimates and their confidence intervals if they were added to the probabilistic model.

Note that many of the uncertainties encountered in a probabilistic exposure assessment also affect deterministic assessments, although differing in detail. However, they are usually not considered explicitly in basic deterministic assessments because the standard conservative assumptions are assumed to provide an appropriate level of protection. In probabilistic assessments, however, it is necessary to consider the uncertainties explicitly, because some of the conservative assumptions of deterministic assessments have been replaced with more realistic distributions.

Note that the evaluations in Tables 7 and 8 are necessarily generic in nature: they do not refer to a specific assessment, but rather to the general range of situations the Panel expects to be encountered by assessors when conducting pessimistic model runs for basic probabilistic assessment. Therefore, *it is recommended that assessors use Tables 7 and 8 as a starting point when performing uncertainty evaluation for specific assessments* that follow the recommended approaches, adjusting the evaluation case by case to take account the specific situation (e.g., amount and quality of data) affecting the assessment in hand. Similarly, the evaluations in Tables 7 and 8 may be used as a starting point when evaluating uncertainties for refined probabilistic assessments.

Table 7: Generic evaluation of uncertainties for the approaches used in the **pessimistic model run** for **basic probabilistic assessments** of **ACUTE** exposure. The +/- symbols represent a subjective assessment of the extent to which each source of uncertainty has the potential to cause over-estimation (+) or under-estimation (-) **of the highest simulated exposures**. The number of symbols provides a subjective evaluation of the magnitude of the effect, as illustrated in the graphical scale below. Where the effect is uncertain, or could vary from assessment to assessment, lower and upper evaluations are given (e.g., - / ++ or + / ++). Finally, the combined impact of all the uncertainties is considered at the end of the table. Key to probabilistic methods: * fixed value; § variability modelled empirically, §§ variability modelled parametrically, ¶ uncertainty modelled by bootstrapping, ¶¶ uncertainty modelled parametrically.



Assessment component	Approach in pessimistic model run	Subjective evaluation of impact on the upper tail exposures	Brief explanation of evaluation
1. Modelling food consumption	Empirical + bootstrap; examine which commodities contribute to upper tail exposures §¶	● (common foods and large survey) - - / ● (small survey and/or rare foods)	Model is limited to intakes observed in survey. With large surveys this will cause no underestimation for common foods. Tendency to underestimation if there is limited data for the foods driving exposure
2. Use of old food consumption survey data	Not considered	● (*sometimes* - /+)	Little effect unless consumption has recently changed for a food with high residues
3. Measurement/reporting uncertainty in consumption surveys.	Not considered	- /+	Uncertainty could be larger for some foods (e.g., rice, potatoes) than others (e.g.fruit eaten whole) (Martine Bakker, pers. comm.).
4. Over-reporting of fruit and vegetable consumption	Not considered	●/+	People might exaggerate intake by up to 2-fold.
5. Relation of consumption to body weight	Kept together in dietary survey records	●	Kept together in dietary survey records
6. Water consumption	Treat as food if included in dietary survey, or use deterministic estimate from drinking water assessment.*	●	Survey may not identify all water consumed, e.g., in recipes. Contribution to exposure generally limited due to low residues in water.
7. Food conversion factors (recipes)	Use available recipe databases*	- /+	Recipes expected to be within +/-x2 for most ingredients/foods. Could be more for minor ingredients, but less likely to drive exposure.

8. Unit weights	Use same values as in deterministic assessments*	- /●	Fixed unit weights are used instead of a distribution. Extreme units might be up to about 2x the default values.
9. Residue definitions	Use residue definition for risk assessment, applying conversion factor where appropriate*. Evaluation of unquantified uncertainties.	- - /++ (?)	Ratios of metabolite to parent vary by up to 50 fold over time in some cases (EFSA Scientific Panel on Plant Protection Products and their Residues (PPR), 2012, Appendix D). Those measured in supervised trials/feeding studies may be smaller than would apply at the time of consumption, although the effect of this on exposure may be reduced if it applies mostly to smaller residues. Also, the toxicity of the metabolite may be lower or higher than parent. Difficult to assess impact of these factors.
10. Residue measurement uncertainty	Not modelled.	+ (large datasets) -/+ (small datasets)	Expected to be up to 50%, see section 4.2.11. For large datasets, will increase variance and hence upper tail exposures. Impact on assessment more uncertain when number of measured residues is small.
11. Unmeasured residues in animal commodities	MRL or default MRL (0.01)*	●/+ (may be +++ for milk)	When MRL is default, may be far above true residues. When based on data, MRL could sometimes underestimate upper tail residues but they are unlikely to more than 2x the MRL. Effect on exposure limited unless milk is a driver of exposure.
12. Between lot/sample variation of residues	Lognormal for positive values (if n>2) §§	- - /++	True distribution of positive sample residues will tend to approximate lognormal, as shown for some examples (Boon et al.)2003b), but is likely to deviate in detail. Very uncertain when n is low, especially in the tails. 'Best fit' lognormal may over- or under-estimate frequency of residues that lead to exposures in the region of the ARfD
13. Sampling uncertainty for lot/sample residues	Parametric model (if >2 positive values) ¶¶	- /++++	The upper confidence limit should generally be protective if the distribution approximates lognormal, and could be very conservative.

14. Treatment of residues below LOR[69]	Set <LOR to LOR*	●/+	Limited effect on level of upper tail acute exposures except when these are driven strongly by foods where the highest measured residues are close to the LOR, or where many foods contribute to individual person–days in the upper tail.
15. Sampling uncertainty of proportion of residues below LOR	Parametric model ¶¶	●/+	Upper confidence limit should generally be protective but increases frequency of positive residues rather than increasing their magnitude, especially where the highest measured residues are close to the LOR, or where many foods contribute to individual person-days in the upper tail.
16. Use of supervised trial residues data for focal commodity in authorisation scenarios, or as a substitute for limited monitoring data	Assume residues at level of supervised trial	- /● (for focal commodity in authorisation scenario) - /++ (in all other cases)	In authorisation scenario, 100% of the focal commodity is assumed to be treated at critical GAP. 0-5% of residues are expected to exceed the MRL (see section 3). When used as a substitute for monitoring data, trial data will overestimate the majority of residues but occasional residues above the MRL are still expected from applications made at the critical GAP.
17. Percent crop treated (when using supervised trials/feeding studies data)	Assume 100% of crop treated*	● (for focal commodity in authorisation scenario) ●/++ (in all other cases)	In authorisation scenario, 100% of the focal commodity is assumed to be treated. In assessments of monitoring data, true % treated always <100%, often much lower. Will over-estimate *frequency* of high exposures 2x-100x. *Level* of high exposures over-estimated to smaller degree due to increased frequency of positive residues in multiple foods in same day.
18. Changes over time in use patterns (e.g., application rates).	Not considered	●	Over time, users might shift more to using the worst case GAP (e.g., due to pest pressure): this would increase the frequency of residues close to MRL but not their levels.

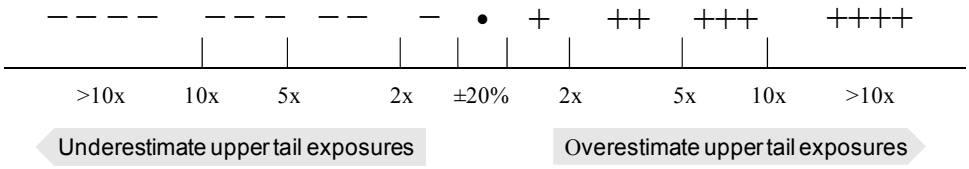
[69] LOR: Limit of Reporting.

19. Limited amounts of monitoring data	Use appropriate data from other countries, other commodities or supervised trials/feeding studies	- - /+++	Frequency and levels of positives may be higher or lower in surrogate data but will not cancel out when upper tail exposures are driven mainly by single foods. Effect could vary from limited under-estimation to large over-estimation (latter when using supervised trials data).
20. No supervised trials (as substitute for monitoring data)	Use appropriate trials data from other commodities	- /+ (more for minor commodities)	If extrapolation guidance is followed, substitute trial residues should be within a factor of 2 for major commodities (e.g., apples and pears), more for minor commodities.
21. Residues for non-authorised use	Treat as for authorised uses except set <LOR to zero*	●	Positives <LOR will make minimal contribution to upper tail exposures.
22. Mean residue of focal commodity in high residue event	Model uncertainty due to sample size and apply unit variability model §§¶¶	●/++	If sampling uncertainty is modelled using a variance from a crop with similar variability, upper confidence bound should usually be conservative. For unit variation see below.
23. Between unit variation (e.g., variability factors, VF, or coefficient of variation CV)	Beta or Lognormal – conservative VF or CV §§	Effect depends on VF/CV used: e.g., if use VF=6.8 ●/++, if use VF=1 - -/●	Effect will depend on value assumed for VF or CV (see section 4.2.9) and will be larger when examining further into the upper tail of exposures.
24. Residues in prepared foods	Assume prepared from the same sample of raw commodity as any raw consumption, and include unit variability §§	●/++	Realistic estimate when food was prepared from raw commodity, but overestimate when the prepared food was purchased.
25. Relation monitoring to residues encountered by consumer	Monitoring data assumed to be representative of residues encountered by consumer	●/+	Residue data from monitoring samples may overestimate the real exposure of the consumer, due to the fact that sampling can be done at several points in the distribution chain (e.g., farm gate, retailer, supermarket) and that at the time of consumption the residue may have declined. However, high exposures likely to be driven by fresh produce, which will usually have a relatively short time from sampling to consumption, so potential reduction is limited.

26. Processing factors	Set to 1 (no change) or use highest individual measured value, whichever is highest*	●/+++	Realistic estimate when processing has little effect. Large over-estimate when processing effect is large. Some underestimation when concentration occurs, due to small number of measurements, but rarely a major driver of exposure.
27. Cumulative assessment – selection of substances to include in the CAG	Assess by repeating assessment including or excluding substances whose membership of CAG is uncertain	● if all potentially relevant substances are examined, otherwise unknown.	If all potentially relevant substances are examined by repeating assessment including and excluding them, then the uncertainty is covered. Otherwise the impact of substances that are not considered is unknown.
28. Cumulative assessment – relative potency factors (RPFs)	Assume slopes of dose-response curves are parallel for all substances in Cumulative Assessment Group	- /+ if slopes close to parallel, possibly more if RPFs based on NOELs. Impact unknown if slopes not parallel.	For some CAGs it has been shown that there is some variation in slope. The impact of this on the assessment requires further investigation.
29. Cumulative assessment – imputation of missing residue data	Simple conservative method for imputation of residues	●/++	Proposed method should be conservative but effect of this is limited in upper tail because this will be driven by combinations of positive residues whereas majority of imputed residues will be <LOR.
30. Residues in water	Assume legal limit (0.1ppb for single substance, 0.5ppb for cumulative assessment)*	●/+	Residues in water sometimes exceed the legal limit. Legal limit usually an overestimate but limited impact on exposure unless this is driven by water.
31. Residues from rotational/succeeding crops	Estimated by expert judgment (see section 4.2.1)	●	Very unlikely to be a driver of high exposures compared to residues deriving from applications to the primary crop except for some metabolites – not very often.
32. Targetted monitoring	May be ignored (see section 4.2.8)	●/++ if targeted samples are included	Including samples from targeted sampling performed by authorities is more likely to affect commodities that have high residues and cause high exposures. Targetting will increase their frequency more than the levels, although increased sampling will reveal some higher residues.

OVERALL ASSESSMENT	The overall impact of all the uncertainties should be evaluated case by case for each probabilistic assessment. First, the assessor should review all the individual uncertainties above, and adjust the evaluations as appropriate to the considerations relevant for their assessment, including the amount and quality of data used. Second, the assessor should consider all the uncertainties together and form a subjective judgement of the overall uncertainty affecting the assessment outcome. This should not be done by any simple summation of the symbols for individual uncertainties, but by using expert judgement to consider the overall impact, taking account of any potential dependencies between the individual uncertainties. The overall conclusion should be expressed using the same symbols and scale as for the individual uncertainties and accompanied by a narrative explanation of the reasoning used by the assessor in reaching their overall judgement.

Table 8: Generic evaluation of uncertainties for the approaches used in the **pessimistic model run** for **basic probabilistic assessments** of **CHRONIC** exposure. The +/- symbols represent a subjective assessment of the extent to which each source of uncertainty has the potential to cause over-estimation (+) or under-estimation (-) **of the highest simulated exposures**. The number of symbols provides a subjective evaluation of the magnitude of the effect, as illustrated in the graphical scale below. Where the effect is uncertain, or could vary from assessment to assessment, lower and upper evaluations are given (e.g., - / ++ or + / ++). Finally, the combined impact of all the uncertainties is considered at the end of the table. Key to probabilistic methods: * fixed value; § variability modelled empirically, §§ variability modelled parametrically, ¶ uncertainty modelled by bootstrapping, ¶¶ uncertainty modelled parametrically.



Assessment component	Approach in pessimistic model run	Subjective evaluation of impact on the upper tail exposures	Brief explanation of evaluation
1. Modelling consumption (modelled as exposure, after combination with residues)	Observed Individual Means method + bootstrap; examine which commodities contribute to upper tail exposures. §¶	- /++++	Impact depends on what part of upper tail is considered. At 95%ile there could in rare cases be underestimation by up to 2x; in extreme tail expect large overestimates.
2. Use of old food consumption survey data	Not considered	• (*sometimes* - /+)	Little effect unless consumption has recently changed for a food with high residues
3. Measurement/reporting uncertainty in consumption surveys.	Not considered	- /+	Uncertainty could be larger for some foods (e.g., rice, potatoes) than others (e.g.fruit eaten whole) (Martine Bakker, pers. comm.).
4. Over-reporting of fruit and vegetable consumption	Not considered	•/+	People might exaggerate intake by up to 2-fold.
5. Relation of consumption to body weight	Kept together in dietary survey records	•	Kept together in dietary survey records
6. Water consumption	Treat as food if included in dietary survey, or use deterministic estimate from drinking water assessment.*	•	Survey may not identify all water consumed, e.g., in recipes. Contribution to exposure generally limited due to low residues in water.
7. Food conversion factors (recipes)	Use available recipe databases*	- /+	Recipes expected to be within +/-x2 for most ingredients/foods. Could be more for minor ingredients, but less likely to drive exposure.

8. Residue definitions	Use residue definition for risk assessment, applying conversion factor where appropriate*. Evaluation of unquantified uncertainties.	- - /++ (?)	Ratios of metabolite to parent vary by up to 50 fold over time in some cases (EFSA Scientific Panel on Plant Protection Products and their Residues (PPR), 2012, Appendix D). Those measured in supervised trials/feeding studies may be smaller than would apply at the time of consumption, although the effect of this on exposure may be reduced if it applies mostly to smaller residues. Also, the toxicity of the metabolite may be lower or higher than parent. Difficult to assess impact of these factors.
9. Residue measurement uncertainty	Not modelled.	●	Expected to be up to 50%, see section 4.2.11. Effect on mean residue will tend to average out except in small datasets.
10. Unmeasured residues in animal commodities	MRL or default MRL (0.01)*	●/+ (may be +++ for milk)	When MRL is default, may be far above true residues. When based on data, MRL could sometimes underestimate upper tail residues but they are unlikely to more than 2x the MRL. Effect on exposure limited unless milk is a driver of exposure.
11. Mean residue for each commodity	Mean of available data*	- /+	Sampling uncertainty could have a substantial effect on median estimate if the number of measurements is small for the foods driving exposure.
12. Sampling uncertainty for mean residues	Empirical bootstrap of available data ¶	●/++++	The upper confidence limit should generally be protective, and could be very conservative.
13. Treatment of residues below LOR[70]	Set <LOR to LOR*	●/++	Effect will be largest when the positive measured values for the foods driving exposure are close to the LOR. Effect will be larger in cumulative assessments involving multiple substances with many foods.
14. Use of supervised trial residues data for focal commodity in authorisation scenarios, or as a substitute for limited monitoring data	Assume residues at level of supervised trial	● (for focal commodity in authorisation scenario) ●/++ (in all other cases)	In authorisation scenario, 100% of the focal commodity is assumed to be treated at critical GAP. When used as a substitute for monitoring data, trial data will overestimate many of the residues encountered by consumers and hence overestimate the mean residue.

[70] LOR: Limit of Reporting.

15. Percent crop treated (when using supervised trials/feeding studies data)	Assume 100% of crop treated*	● (for focal commodity in authorisation scenario) ●/++++ (in all other cases)	In authorisation scenario, 100% of the focal commodity is assumed to be treated. In assessments of monitoring data, true % treated can often be less than 10% which will reduce mean residues and hence long term exposures by over 10x.
16. Changes over time in use patterns (e.g., application rates).	Not considered	●	Over time, users might shift more to using the worst case GAP (e.g., due to resistance or pest pressure): this would increase the frequency of residues close to MRL but not their levels.
17. Limited amounts of monitoring data	Use appropriate data from other countries, other commodities or supervised trials/feeding studies	- - /+++	Frequency and levels of positives may be higher or lower in surrogate data but will not cancel out when upper tail exposures are driven mainly by single foods. Effect could vary from limited under-estimation to large over-estimation (latter when using trials data).
18. No supervised trials (as substitute for monitoring data)	Use appropriate trials data from other commodities	- /+ (more for minor commodities)	If extrapolation guidance is followed, substitute trial residues should be within a factor of 2 for major commodities (e.g., apples and pears), more for minor commodities.
19. Residues for non-authorised use	Treat as for authorised uses except set <LOR to zero*	●	Positives <LOR will make minimal contribution to upper tail exposures.
20. Relation monitoring to residues encountered by consumer	Monitoring data assumed to be representative of residues encountered by consumer	●/+	Residue data from monitoring samples may overestimate the real exposure of the consumer, due to the fact that sampling can be done at several points in the distribution chain (e.g., farm gate, retailer, supermarket) and that at the time of consumption the residue may have declined. However, high exposures likely to be driven by fresh produce, which will usually have a relatively short time from sampling to consumption, so potential reduction is limited.
21. Processing factors	Distribution of estimates for mean processing factor, obtained by bootstrapping measured values. *	●/+++	Realistic estimate when processing has little effect. Large over-estimate when processing effect is large. Some underestimation of concentration due to small number of measurements.

22. Cumulative assessment – selection of substances to include in the CAG	Assess by repeating assessment including or excluding substances whose membership of CAG is uncertain	● if all potentially relevant substances are examined, otherwise unknown.	If all potentially relevant substances are examined by repeating assessment including and excluding them, then the uncertainty is covered. Otherwise the impact of substances that are not considered is unknown.
23. Cumulative assessment – relative potency factors	Assume slopes of dose-response curves are parallel for all substances in Cumulative Action Group	- /+ if slopes close to parallel, otherwise unknown	For some CAGs it has been shown that there is some variation in slope. The impact of this on the assessment requires further investigation.
24. Cumulative assessment – imputation of missing residue data	Simple conservative method for imputation of residues	●/++	Proposed method should be conservative but effect of this is limited in upper tail because this will be driven by combinations of positive residues whereas majority of imputed residues will be <LOR.
25. Residues in water	Assume legal limit (0.1ppb for single substance, 0.5ppb for cumulative assessment)*	●/++	Residues in water sometimes exceed the legal limit, but average level in water is likely to be less than half legal limit.
26. Residues from rotational/succeeding crops	Estimated by expert judgment (see section 4.2.1)	●	Will only be relevant for a proportion of applications. Limited impact on mean residues.
27. Targetted monitoring	May be ignored (see section 4.2.8)	+/●	Including samples from targeted sampling performed by authorities is more likely to affect commodities that have high residues and cause high exposures. Targetting will increase their frequency more than the levels, although increased sampling will reveal some higher residues.
OVERALL ASSESSMENT	The overall impact of all the uncertainties should be evaluated case by case for each probabilistic assessment. First, the assessor should review all the individual uncertainties above, and adjust the evaluations as appropriate to the considerations relevant for their assessment, including the amount and quality of data used. Second, the assessor should consider all the uncertainties together and form a subjective judgement of the overall uncertainty affecting the assessment outcome. This should not be done by any simple summation of the symbols for individual uncertainties, but by using expert judgement to consider the overall impact, taking account of any potential dependencies between the individual uncertainties. The overall conclusion should be expressed using the same symbols and scale as for the individual uncertainties and accompanied by a narrative explanation of the reasoning used by the assessor in reaching their overall judgement.		

GLOSSARY AND ABBREVIATIONS

Acute exposure
A contact between an agent and a target occurring over a short time, generally less than a day. (Other terms, such as "short-term exposure" and "single dose" are also used (ISEA, 2005).

Acute reference dose (ARfD)
Estimate of the amount of a substance in food and/or drinking water, normally expressed on a body weight basis, that can be ingested in a period of 24 h or less without appreciable health risk to the consumer on the basis of all known facts at the time of the evaluation (JMPR).

Acute toxicity
Adverse effects of finite duration occurring within a short time (up to 14 d) after administration of a single dose (or exposure to a given concentration) of a test substance or after multiple doses (exposures), usually within 24 h of a starting point (which may be exposure to the toxicant, or loss of reserve capacity, or development change, etc. (IUPAC, 2006).

Background commodity
A commodity other than the focal commodity (one for which an MRL is to be set or for which a high residue event has been monitored), but which could contribute to the overall exposure.

Bayesian
The Bayesian or subjective view is that the probability of an event is the degree of belief that a person has, given some state of knowledge, that the event will occur. In the classical or frequentist view, the probability of an event is the frequency with which an event occurs given a long sequence of identical and independent trials. In exposure assessment situations, directly representative and complete data sets are rarely available; inferences in these situations are inherently subjective. The decision as to the appropriateness of either approach (Bayesian or Classical) is based on the available data and the extent of subjectivity deemed appropriate (U.S EPA, 1997a).

Benchmark dose
BMD
A dose or concentration that produces a predetermined change in response rate of an adverse effect (called the benchmark response or BMR) compared to background (EFSA, 2008).

Bimodal distribution
A continuous probability distribution with two different modes. These appear as distinct peaks (local maxima) in the probability density function.

Bootstrapping
Random resampling technique (Efron 1993, U.S EPA 1997a).

Censored data
See Left-censored data.

Chronic exposure
A continuous or intermittent long-term contact between an agent and a target. Other terms, such as "long-term exposure," are also used. (ISEA, 2005).

Chronic effect
Consequence that develops slowly and/or has a long lasting course: may be applied to an effect that develops rapidly and is long-lasting (IUPAC, 2006).

Chronic toxicity
1. Adverse effects following chronic exposure. 2. Effects that persist over a long period of time whether or not they occur immediately upon exposure or are delayed (IUPAC, 2006).

Commodity
A human food or animal feed crop (e.g., carrots, apples, corn, eggs or orange juice).

Conversion factor:
Food conversion factor: Multiplicative factor used to convert food products as eaten and recorded in dietary surveys to the corresponding raw agricultural commodities, or these commodities in the forms for which monitoring data are available. **Residue conversion factor:** multiplicative factor applied to monitoring data in order to take into account the exposure to metabolites that are not measured during the monitoring; used when the residue definition for monitoring and risk assessment differ, but address the same toxicological end-point.

Composite sample
A sample formed by combining multiple units of the same commodity for analysis. Formally referred to as "laboratory sample" in Directive 96-23 on Official Control of Pesticide Residues (defined there as "The sample sent to, or received by, the laboratory. A representative quantity of material removed from the bulk.")

Consumption
Food consumption data reflects what either individuals or groups consume in terms of solid foods, beverages, including drinking water, and supplements. Food consumption can be estimated through food consumption surveys at an individual (Individual dietary surveys) or household level (Household budget surveys) or approximated through food supply data derived from food balance sheets.

Cumulative Assessment Group (CAG)
A group of chemicals that could plausibly act by a common mode of action, not all of which will necessarily do so. Membership of a CAG can usually be refined (reduced) by application of successively higher tiers of assessment (EFSA, 2008).

Cumulative exposure assessment
An exposure assessment which considers potential human health risks from all pathways of dietary and nondietary exposures to more than one pesticide acting through a common mechanism of toxicity. When limited to multichemical assessment through one pathway (e.g., dietary), this may be called a cumulative dietary exposure assessment.

Exposure
Contact between an agent and a target. Contact takes place at an exposure surface over an exposure period (ISEA, 2005).

Exposure Assessment
The qualitative and /or quantitative evaluation of the likely intake of biological, chemical or physical agents via food as well as exposure from other sources if relevant (WHO, 1995).

Empirical modelling
Modelling based on empirical observations rather than on mathematically-described relationships of the system modeled.

Feeding study
Feeding studies are supervised studies where farm animals are receiving controlled oral doses of unlabelled compounds (incorporated into the feed or applied in another suitable form, e.g.,

in capsules) to establish the relationship between residue levels in feed and likely residues in tissues, milk and eggs.

Focal commodity
A commodity for which an MRL is to be set or for which a high residue event has been monitored, and which is therefore the focus of an exposure assessment.

GAP, Good Agricultural Practice
GAP means the nationally recommended, authorised or registered safe use of plant protection products under actual conditions at any stage of production, being storage, transport, distribution and processing of food and feed. It also implies the application, in conformity with Directive 91/414/EEC, of the principles of integrated pest control in a given climate zone, as well as using the minimum quantity of pesticides and setting MRLs/temporary MRLs at the lowest level which allows the desired effect to be obtained (MRL Regulation).

High residue event
A term used in this document to refer to the finding of a lot of commodity in the marketplace with one or more measured residue values exceeding the MRL.

Highest residue (HR)
The HR is the highest residue level (expressed as mg/kg) in a composite sample of the edible portion of a food commodity when a pesticide has been used according to maximum GAP conditions. The HR is estimated as the highest of the residue values (one from each trial) from supervised trials conducted according to maximum GAP conditions, and includes residue components defined by the JMPR for estimation of dietary intake.

Import tolerance
Defined by Regulation 396/2005 as an MRL set for imported products to meet the needs of international trade where:
— the use of the active substance in a plant protection product on a given product is not authorised in the Community for reasons other than public health reasons for the specific product and specific use; or
— a different level is appropriate because the existing Community MRL was set for reasons other than public health reasons for the specific product and specific use.

Index compound (IC)
Under the RPF approach, one member of the CAG is selected as the index compound which is used as the point of reference for standardizing the potency of other members of the CAG.

Observed Individual Means approach (OIM)
An approach for estimating longer term exposures by taking each individual's observed mean consumption over the duration of a dietary survey.

Left-censored data
Low measured levels of pesticide residues for which an accurate value is not available, because these levels have been reported as being below a Limit of Reporting (LOR).

Limit of reporting (LOR)
A lower limit of residue concentration below which measured levels are not reported. Note that the definition used here is different from the Reporting Limit (RL) as defined by SANCO (2011a). The term LOR encompasses other limits that may be included in datasets used for probabilistic modelling (e.g., LOD, LOQ) but are not distinguished in this document as all are censored values and treated in the same way in the approaches proposed here.

Maximum residue limit (MRL)
Maximum concentration of a residue that is legally permitted or recognized as acceptable in, or on, a food, agricultural commodity, or animal feedstuff as set by Codex or a national regulatory authority (IUPAC, 2006).

Monte Carlo analysis
Monte Carlo analysis is a computer-based method of analysis that uses statistical sampling techniques in obtaining a probabilistic approximation to the solution of a mathematical equation or model (U.S EPA, 1997a).

Margin of exposure (MOE)
Ratio of a toxicological reference dose to estimated exposure.

Monitoring data
In this document, 'monitoring data' refers to data on pesticide residues in food occurring as a result of commercial use, obtained by analysis of samples taken at relevant points in the food chain from producer to marketplace. In other contexts, residues may be monitored in other media, e.g., soil, water, air, etc..

Pesticide residue
Substance which remains in or on food commodity, soil, air, or water following use of a pesticide. For regulatory purposes, it includes the parent compound and any specified derivates such as degradation and conversion products, metabolites, and impurities considered to be of toxicological significance (44, FAO) (IUPAC, 2006).

Percent crop treated
Percentage of a raw agricultural commodity, in the marketplace that is relevant to an exposure assessment, that has been treated with the pesticide under assessment.

Processing factor
Residue level of a specific pesticide in the processed products divided by the residues level in the starting commodity, usually raw agricultural commodity (RAC). Processing factor = residue level (mg kg^{-1}) in processed product/residue level (mg kg^{-1}) in RAC. *Note: Alternative terms sometimes used for processing factor are "concentration factor" when residue levels increase and "reduction factor" (inverse of processing factor) when residue levels decrease* (IUPAC, 2006).

Processed food
Product resulting from the application of physical, chemical, or biological processes, or combinations of these (e.g., canning), to a primary food commodity, and intended for sale to the consumer, for use as an ingredient in the manufacture of a food product or for further processing. (IUPAC, 2006).

Random Variable
A random variable is a quantity which can take on any number of values but whose exact value cannot be known before a direct observation is made. For example, the outcome of the toss of a pair of dice is a random variable, as is the height or weight of a person selected at random from the New York City phone book. (U.S EPA, 1997a).

Raw agricultural commodity (RAC)
An unprocessed human food or animal feed crop (e.g., raw carrots, apples, corn or eggs) (US EPA, 1997b).

Relative Potency Factor (RPF)
The ratio of the toxic potency of a given chemical to that of an index chemical in the Cumulative Assessment Group (CAG). Relative potency factors are used to convert exposures of all chemicals in the CAG into their exposure equivalents of the index chemical (EFSA, 2008).

Resample
Drawing repeated samples; in the context of this document, drawing samples of the same size randomly with replacement from a single original dataset.

Uncertainty
Uncertainty refers to lack of knowledge about specific factors, parameters, or models. For example, we may be uncertain about the mean concentration of a specific pollutant at a contaminated site or we may be uncertain about a specific measure of uptake (e.g., 95th percentile fish consumption rate among all adult males in the United States). Uncertainty includes parameter uncertainty (measurement errors, sampling errors, systematic errors), model uncertainty (uncertainty due to necessary simplification of real-world processes, mis-specification of the model structure, model misuse, use of inappropriate surrogate variables), and scenario uncertainty (descriptive errors, aggregation errors, errors in professional judgment, incomplete analysis) (U.S EPA, 1997a).

Variability
Variability refers to observed differences attributable to true heterogeneity or diversity in a population or exposure parameter. Sources of variability are the result of natural random processes and stem from environmental, lifestyle, and genetic differences among humans. Examples include human physiological variation (e.g., natural variation in bodyweight, height, breathing rates, drinking water intake rates), weather variability, variation in soil types and differences in contaminant concentrations in the environment. Variability is usually not reducible by further measurement or study (but can be better characterized) (U.S EPA, 1997a).

Variability factor
The ratio between the 97.5th percentile and mean of a distribution of unit residues.

Supervised trial
Scientific studies for estimating maximum residue limits in which pesticides are applied to crops or animals according to specified conditions intended to reflect commercial practice after which harvested crops or tissues of slaughtered animals are analyzed for pesticide residues. Usually specified conditions are those which approximate existing or proposed good agricultural practice (EFSA 2008, IUPAC 2006).

Supervised trials median residue (STMR)
The median of the residue value (one from each trial) from supervised trials conducted according to maximum good agricultural practice. (IUPAC, 2006).

Unit weight
In the EU, to quantify the potential acute exposure to pesticide residues, typical unit weights of the single food commodities are necessary for those commodities weighting more than 25 g but lower than 250 g. According to WHO, the unit weight refers to weight of the edible portion of a single unit commodity expressed as mean or median value (EFSA, 2007).