SCIENTIFIC REPORT

Cumulative dietary exposure assessment of pesticides that have acute effects on the nervous system using SAS® software

European Food Safety Authority (EFSA),
Bruno Dujardin and Valentina Bocca

Abstract

Retrospective dietary exposure assessments were conducted for two groups of pesticides that have acute effects on the nervous system: brain and/or erythrocyte acetylcholinesterase inhibition, and functional alterations of the motor division. The pesticides considered in this assessment were identified and characterised in the scientific report on the establishment of cumulative assessment groups of pesticides for their effects on the nervous system. The exposure assessments used monitoring data collected by Member States under their official pesticide monitoring programmes in 2014, 2015 and 2016 and individual food consumption data from 10 populations of consumers from different countries and from different age groups. Exposure estimates were obtained for each group of pesticides by means of a 2-dimensional Monte Carlo simulation, which was implemented in SAS® software. Results were validated against exposure estimates obtained by the Dutch National Institute for Public Health and the Environment (RIVM) using the Monte Carlo Risk Assessment (MCRA) software. Both tools produced nearly identical results and minor differences were mainly attributed to the random effect of probabilistic modelling. The exposure estimates obtained in this report are used in the final scientific report on the cumulative dietary risk characterisation of pesticides that have acute effects on the nervous system. The latter combines the hazard assessment and exposure assessment into a consolidated risk characterisation including all related uncertainties.

The EFSA Journal is a publication of the European Food Safety Authority, an agency of the European Union.

Summary

Retrospective dietary exposure assessments were conducted for two groups of pesticides that have acute effects on the nervous system: brain and/or erythrocyte acetylcholinesterase inhibition, and functional alterations of the motor division. The pesticides considered in this assessment were identified and characterised in the scientific report on the establishment of cumulative assessment groups of pesticides for their effects on the nervous system.

The exposure calculations used monitoring data collected by Member States under their official pesticide monitoring programmes in 2014, 2015 and 2016 and individual food consumption data from 10 populations of consumers from different countries and from different age groups. Regarding the selection of relevant food commodities, the assessment included water, foods for infants and young children and 30 raw primary commodities of plant origin that are widely consumed within Europe.

Exposure estimates were obtained with SAS® software using a 2-dimensional Monte Carlo simulation, which is composed of an inner-loop execution and an outer-loop execution. Variability within the population is modelled through the inner-loop execution and is expressed as a percentile of the exposure distribution. The outer-loop execution is used to derive 95% confidence intervals around those percentiles (reflecting the sampling uncertainty of the input data).

As agreed by risk managers in the Standing Committee on Plants, Animals, Food and Feed (SC PAFF), this methodology is applied in a tiered approach. While the first-tier calculations (Tier I) use very conservative assumptions for an efficient screening of the exposure with low risk for underestimation, the second-tier assessment (Tier II) includes assumptions that are more refined but still intended to be conservative. For each scenario, exposure estimates were obtained for different percentiles of the exposure distribution and the total margin of exposure (MOET, i.e. the ratio of the toxicological reference dose to the estimated exposure) was calculated at each percentile. In accordance with the threshold agreed at the SC PAFF, further regulatory consideration would be required when the MOET calculated at the 99.9th percentile of the exposure distribution is below 100.

The lowest MOET estimates were obtained for pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition. According to the Tier II scenario, MOET estimates at the 50th, 95th and 99th percentile of the exposure distribution were all well above 100. At the 99.9th percentile, estimates were mostly below 100, ranging from 40.2 to 61.7 in toddlers and other children. For adults, the MOETs were ranging from 92.4 to 121. The exposure to this group of pesticides was predominantly driven by the occurrence of chlorpyrifos, triazophos and omethoate. Other important contributors included dichlorvos, formetanate and carbofuran.

For pesticides associated with functional alterations of the motor division, MOETs calculated at the 99.9th percentile of the exposure distribution were higher, ranging from 66.3 to 89.5 in toddlers and other children and from 145 to 181 in adults. In this case, the main drivers for the exposure were identified as triazophos and deltamethrin. Other important contributions came from beta-cypermethrin, omethoate, thiram, chlormequat and acrinathrin.

To ensure a rigorous validation of the methodology, exposure estimates obtained by the European Food Safety Authority (EFSA) were validated against those obtained by the Dutch National Institute for Public Health and Environment (RIVM) using the Monte Carlo Risk Assessment (MCRA) software, version 8.3. Comparison of the results revealed that both tools produced nearly identical results and any observed differences are mainly attributed to the random effect of probabilistic modelling. These minor differences, however, do not impact on the outcome of the exposure assessment.

It is important to note that all calculations were conducted with conservative assumptions likely to overestimate the exposure, even in the more refined Tier II scenario. The most important overestimation originates from the limited availability of processing factor data. When such data are missing, it is assumed that all pesticides in the raw primary commodity will reach the end consumer without any loss of residues. Sensitivity tests have demonstrated that, for pesticides associated with functional alterations of the motor division, further data on the effect of processing might result in a threefold increase of MOET estimates in toddlers. Information on the effect of processing prior to consumption is therefore expected to substantially reduce the uncertainty, particularly in the case of risk drivers.

Another impactful assumption is the random assignment of active substances to unspecific residue definitions (i.e. a measurement that may comprise multiple active substances). All substances were assumed to have an equal probability of being applied while this is not necessarily reflecting the actual use of the pesticides. Comparison with the Tier I scenario has indicated that this assumption is very impactful. If data on the use frequency of pesticides would be available at the European Union (EU) level, this assumption could be further refined.

Uncertainties caused by the imputation of left-censored data (i.e. measurements below the limit of quantification), were found to have a very limited impact. This finding is consistent with the expectation that MOETs at the 99.9th percentile of the acute exposure distribution are primarily driven by samples with quantifiable findings.

Analysis of the factors influencing the upper part of the exposure distribution also revealed that the highest acute exposures do not result from cumulative exposure to multiple compounds. For toddlers and children, around 75% of the high exposure estimates were driven by a single substance in a specific commodity. For pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition, around 95% of these combinations were found to exceed the maximum residue level (MRL) that was in place at the end of the reference period. It should be noted, however, that MRLs for some of these combinations were lowered shortly before the end of the reference period (e.g. MRLs for chlorpyrifos in apples and pears were lowered in August 2016). It is therefore expected that, in the meantime, exposure from these combinations will have decreased already.

Uncertainties considered in this assessment, however, only refer to the exposure calculations and should still be considered in conjunction with other uncertainties that may apply to the hazard characterisation. Hence, together with the results obtained by RIVM, the exposure estimates presented in this report are used for the final scientific report on the cumulative dietary risk characterisation of pesticides that have acute effects on the nervous system. The latter combines the hazard assessment and exposure assessment into a consolidated risk characterisation including all related uncertainties.

Table of contents

1.   Introduction

Regulation (EC) No 396/2005 establishes the rules governing the setting of maximum residue levels (MRLs) for pesticides at European level. Article 14 of that Regulation stipulates that decisions on applications concerning MRLs shall take into account known cumulative and synergistic effects of pesticides when the methods to assess such effects are available. Likewise, Article 4 of Regulation (EC) No 1107/2009 further requires that the residues of the plant protection products shall not have any harmful effects on human health or animal health, taking into account known cumulative and synergistic effects where the scientific methods accepted by the Authority to assess such effects are available.

To support the implementation of cumulative risk assessment to pesticide residues, EFSA's Scientific Panel on Plant Protection Products and their Residues (PPR Panel) adopted two scientific opinions regarding the development of a tiered methodology for cumulative risk assessment to pesticide residues (EFSA PPR Panel, 2008, 2009) and two scientific opinions regarding the methodology to establish cumulative assessment groups (CAGs) for pesticides (EFSA PPR Panel, 2013a,b). A guidance document on the use of a probabilistic methodology for the dietary exposure assessment of pesticides, including cumulative exposure, was also adopted (EFSA PPR Panel, 2012). Aside from the general methodological principles, this guidance document also describes different scenarios for probabilistic modelling. First two basic modelling scenarios were proposed, i.e. optimistic and pessimistic. While the optimistic scenario is aimed at calculating the upper estimates of the true distribution of exposure, the pessimistic scenario is intended to obtain the lower estimates of that distribution. The outcome of both scenarios can then be used to determine whether further refinement of the exposure assessment is necessary.

Meanwhile, the European Commission funded the ACROPOLIS project under its 7th Framework Programme for Research (FP7). The main outcome of this project, coordinated by the Dutch National Institute for Public Health and Environment (RIVM), was the release of the Monte Carlo Risk Assessment (MCRA) software, version 8. This is a web-based software that allows performing higher tier exposure assessment to multiple pesticides. This software complies with the EFSA guidance document on the use of probabilistic methodologies for dietary exposure assessment to pesticides.

During the Standing Committee on Plants, Animals, Food and Feed (SC PAFF) of 11–12 June 2015, a discussion on related risk management aspects took place between the European Commission and Member States. The Standing Committee agreed on several parameters and assumptions that should be applied when assessing cumulative exposure to pesticide residues. The conclusion also included a change of approach, where the basic and refined modelling principles, as initially proposed by the European Food Safety Authority (EFSA) in its guidance document, were replaced by a tiered methodology (Tier I and Tier II).

EFSA therefore decided to consolidate the cooperation with RIVM by means of a Framework Partnership Agreement, which mainly aimed at testing and improving MCRA Software in view of its implementation in cumulative risk assessment to pesticide residues. In particular, the MCRA software was made scalable in order to handle large CAGs (van de Voet et al., 2016) and a proposal for a data model compliant with the MCRA software was also elaborated (Kruisselbrink et al., 2018). A pilot assessment of acute cumulative exposure to pesticide residues was then performed with the adapted MCRA software, version 8.2. This version of MCRA integrated the tiered methodology as agreed by Member States.

For these pilot calculations, RIVM relied on the preliminary CAGs that were available for the nervous system at that time (EFSA PPR Panel, 2013a). These CAGs were further refined and recently finalised by EFSA following a public consultation (EFSA, 2019a). The exposure calculations that used these preliminary data should therefore be revised accordingly. Furthermore, during the SC PAFF of 18–19 September 2018, Member States agreed on additional parameters and assumptions for the assessment of cumulative exposure to pesticide residues. Therefore, the European Commission, in its letter of 10 October 2018,[1] asked EFSA to ensure that the newly agreed parameters would be used when calculating cumulative exposure to pesticides before publishing the assessments.

Due to these additional considerations, under a second Framework Partnership Agreement, EFSA and RIVM agreed to revise the MCRA software in order to handle the revised input data and integrate new functionalities addressing the assumptions and parameters agreed by Member States. RIVM was also requested by EFSA to revise the cumulative exposure assessment for pesticides affecting the nervous system using the new MCRA software (version 8.3). At the same time, EFSA initiated the process of repeating the cumulative exposure assessments internally using a different software that was available

---

[1] http://www.efsa.europa.eu/sites/default/files/SANTE_CRA_Mandate.pdf

in-house, i.e. SAS® software. By comparing results obtained with MCRA software and SAS® software, RIVM and EFSA will ensure a rigorous validation of the methodologies used for exposure assessment.

The current report presents the assessment of acute cumulative exposure to pesticides affecting the nervous system using SAS® software. For these calculations, EFSA applied the main methodological principles described in the guidance document on probabilistic modelling of dietary exposure to pesticide residues (EFSA PPR Panel, 2012). Regarding the scenarios and assumptions, however, instead of using the basic and refined modelling described in the guidance document, EFSA relied on the tiered approach as agreed by Member States in 2018.

Considering that the implementation of cumulative exposure assessment to pesticide residues is still in a pilot phase, it was decided to start with retrospective assessments for the reference period 2014-2016. Retrospective assessments refer to the post-authorisation period of active substances and usually rely on the review of monitoring data. Retrospective assessments are therefore expected to better reflect actual exposure of consumers. Considering the complexity and resources required for this type of assessment, the calculations are limited to 10 dietary surveys that cover different age classes and geographical areas. Likewise, EFSA only considers the food consumption data for water, foods for infants and young children and 30 raw primary commodities (RPCs) of plant origin that are widely consumed within Europe. To ensure consistency with the exposure calculations carried out by RIVM using MCRA software, EFSA uses the same input data as those provided to RIVM.

Together with the reports prepared by RIVM, the results of this report will be used in the final scientific report on the cumulative dietary risk characterisation of pesticides that have acute effects on the nervous system (EFSA, 2019b). The final report will assess all sources of uncertainty identified for either the exposure assessment or the establishment of the CAGs. All uncertainties will then be incorporated into a consolidated risk characterisation. Hence, the present report on exposure assessment will not consider the overall assessment of uncertainties or risks.

2. Data and methodologies

2.1. General principles

The cumulative exposure to pesticide residues was assessed in accordance with the guidance on probabilistic modelling of dietary exposure to pesticide residues (EFSA PPR Panel, 2012). Exposure estimates were obtained using a 2-dimensional method where variability is modelled by means of an inner-loop execution, and uncertainty is modelled through an outer-loop execution (see Figure 1).

The **primary input data** required for modelling cumulative exposure to pesticide residues are occurrence data (i.e. the amounts of pesticide residue that are present in foods) and food consumption data (i.e. the types and amounts of those food consumed in a person's diet). These data are stored in the EFSA Data Warehouse. When the exposure calculations are initiated, the data for the relevant food commodities, active substances and dietary surveys are extracted.

Within the **inner-loop execution**, occurrence data are subject to several simulations and imputations. These adjustments are intended to account for inaccuracies and missing information in the occurrence data set (e.g. unspecific measurements, measurements below the analytical limit of quantification, etc.). The consumption data and adjusted occurrence data are then used to estimate acute dietary exposures using an empirical Monte Carlo simulation (i.e. with 100,000 iterations). This results in a distribution that represents the variability of acute exposures within the population.

The different simulations performed during the inner-loop execution require the use of additional data, referred to as **secondary input data**. This includes various types of data which can be used either for the adjustment of the occurrence data (e.g. authorisation status of the active substance) or for improvement of the exposure estimates (e.g. processing factors).

In order to quantify the uncertainties, the model uses an **outer-loop execution** where the inner-loop execution is repeated several times. Prior to each execution, the original consumption and occurrence data sets are modified by means of bootstrapping, a random resampling technique for quantifying sampling uncertainty. By repeating the inner-loop execution multiple times (i.e. 100), the model produces multiple distributions of exposure. The differences between those distributions reflect the uncertainty around the true distribution of exposures.

During the **output preparation**, summary statistics (i.e. percentiles of exposure) are generated for the multiple distributions, resulting in multiple estimates for each percentile of exposure. From these multiple estimates, confidence intervals around each percentile are produced. Subsequently, in order to identify risk drivers, details on the highly exposed consumers are extracted (i.e. consumers

with exposure exceeding the 99th percentile) and average contributions per food commodity and active substance are calculated.

According to the risk management principles agreed among Member States,[1] the methodology described above is applied in a tiered approach. While the first-tier calculations (Tier I) use very conservative assumptions, the second-tier assessment (Tier II) includes assumptions that are more refined but still intended to be conservative. Furthermore, in order to better understand the impact related to some of the assumptions and uncertainties, several sensitivity analyses were carried out.

All extractions, simulations, imputations and calculations described in the subsequent sections were programmed with SAS® Enterprise Guide 7.1 and SAS® Studio 3.71 (Enterprise Edition).
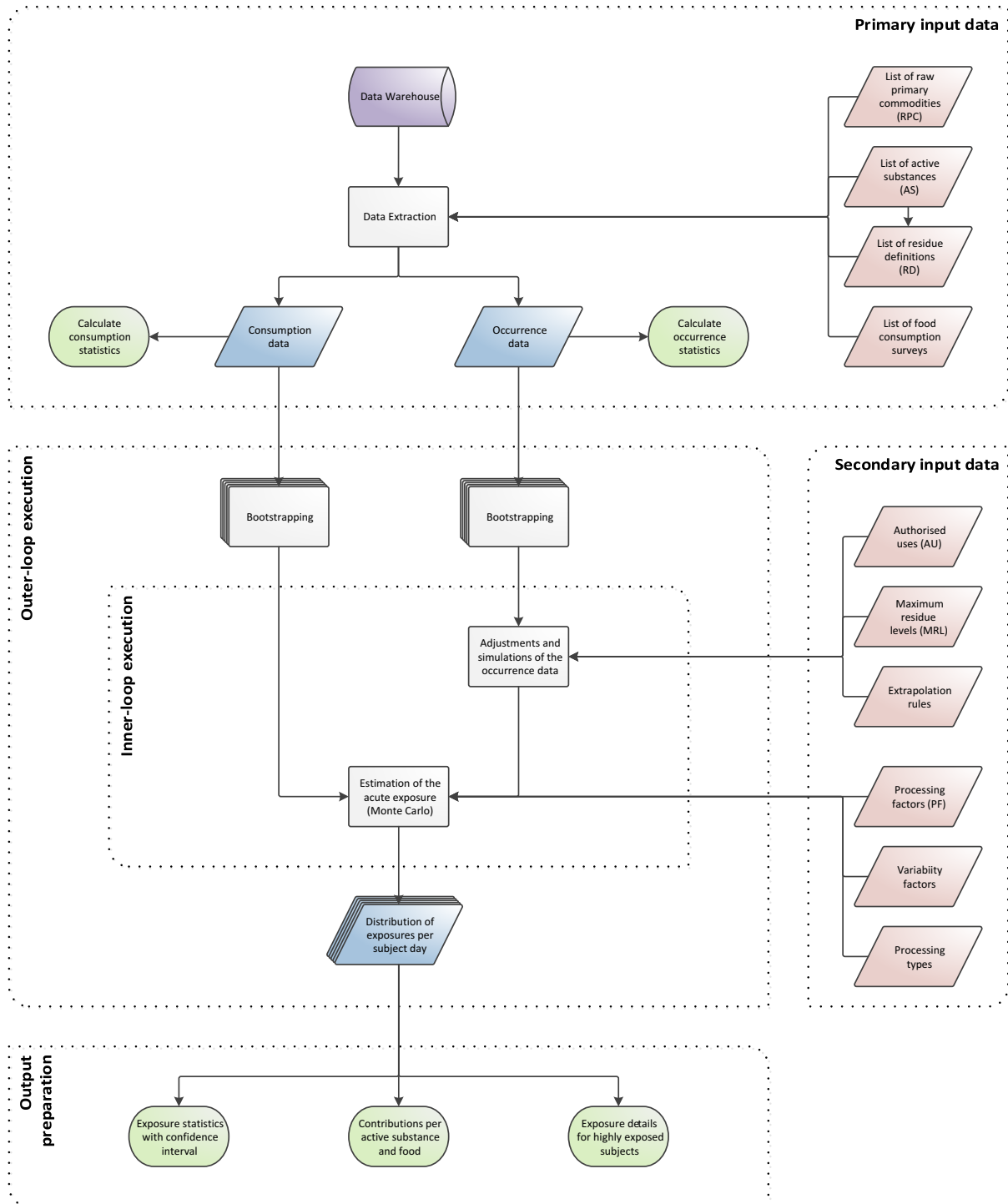

Figure 1: General process for calculating acute cumulative exposure to pesticides

2.2. Primary input data

2.2.1. Raw Primary commodities

To pilot the cumulative exposure assessment to pesticide residues, EFSA selected 30 RPCs of plant origin that are widely consumed in Europe (EFSA, 2015a). Water and foods specifically intended for infants and young children were integrated in the exposure assessment based on their importance in (certain) diets. The full list of the incorporated food commodities is provided in Annex A.1, Table A.1.02 and Annex A.2, Table A.2.02. The variables used to describe each food commodity are reported in Table 1.

For the dietary surveys used in this assessment (see Section 2.2.5), the average contribution of the 30 RPCs to the total consumption of plant commodities (excluding sugar plants) ranges from 73% to 86%. Sugar plants and commodities of animal origin were not considered. As the occurrence of pesticide residues in these commodities is less frequent and at lower levels, their contribution to the dietary exposure is expected to be much lower than the contribution of plant commodities (EFSA, 2019b).

Table 1: Description of the key variables used to describe the raw primary commodities

Name	Label	Description
prodCode	RPC code	Code of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodName	RPC name	Name of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)

RPC: raw primary commodity.

2.2.2. Active substances

Two CAGs were selected by EFSA to pilot the acute cumulative risk assessment to pesticide residues. These CAGs include:

- 47 pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition (CAG-NAN);
- 100 pesticides associated with functional alterations of the motor division (CAG-NAM).

The methodology that was used to identify pesticide active substances affecting the nervous system and the criteria that were applied to define the different assessment groups are described in a separate scientific report (EFSA, 2019a).

For each CAG, the list of active substances, which incorporates the key input data for cumulative exposure assessment, is presented in Annex A.1, Table A.1.01 and Annex A.2, Table A.2.01. The variables contained in the list of active substances are described in Table 2.

The toxicological potency of the active substances within each CAG is defined by means of the no observed adverse effect level (NOAEL). When an index compound is identified for the CAG, toxicological potency may also be expressed as a relative potency factor, i.e. the ratio of the NOAEL to that of the index compound (EFSA PPR Panel, 2012). This allows for the expression of exposure estimates in equivalents of the index compound. In this assessment, however, exposure estimates were normalised to a dimensionless number, referred to as the normalised exposure (see Section 2.4.2). Index compounds and relative potency factors were therefore no longer considered.

Table 2: Description of the variables contained in the list of active substances

Name	Label	Description
paramCode_AS	Substance code	Code of the active substance as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
paramName_AS	Substance name	Name of the active substance as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
NOAEL	NOAEL	No observed adverse effect level (NOAEL) of the active substance (EFSA, 2019a)
Mechanism	Mechanism of action	Short reference to the mechanism of action or to the mode of action, where available (EFSA, 2019a)

Name	Label	Description
Study_type	Study type	Type of regulatory toxicity study required by Regulation (EC) No 1107/2009 from which the NOAEL has been derived (EFSA, 2019a)

AS: active substance.

2.2.3. Residue definitions

While the CAGs are defined at the level of the pesticide active substances, the occurrence data reported to EFSA refer to a residue definition for enforcement purposes (see Section 2.2.4). As the residue definitions, defined by Regulation (EC) No 396/2005, may change over time, single active substances may be associated to multiple residue definitions throughout the reference period. EFSA therefore collected all the residue definitions that were applicable to the selected food commodities and active substances during the reference period 2014–2016. The residue definitions collected for CAG-NAN and CAG-NAM are presented in Annex A.1, Table A.1.03 and Annex A.2, Table A.2.03, respectively.

Depending on metabolism and availability of analytical methods, residue definitions may either be equal to the active substance, include additional metabolites, or even incorporate multiple active substances. When the residue definition includes additional metabolites, which are specific to the active substance, the residue definition is assigned to the active substance assuming that the metabolite will have the same toxicological potency as the parent compound (e.g. sum fipronil and its sulfone metabolite, expressed as fipronil). When the residue definition includes multiple active substances, however, the active substances may have different toxicological potencies (e.g. dithiocarbamates). The latter are referred to as unspecific residue definitions.

When active substances are associated to an unspecific residue definition (e.g. sum of dimethoate and omethoate, expressed as dimethoate), further distinction is made between exclusive and non-exclusive associations.

- Supposing that omethoate would be applied to the field, omethoate cannot be metabolised into dimethoate and the measured residue would be attributed to omethoate only. In this case, the association is considered exclusive.
- Supposing that dimethoate would be applied to the field, dimethoate would partially metabolise into omethoate. In this case, only a proportion of the measured residue would be attributed to dimethoate and the remaining part would be attributed to omethoate. Hence, the association is not exclusive.

Data on the proportions, however, were not readily available to EFSA. Therefore, a default proportion of 0.5 ($\approx$ 50%) was assumed for all associations that are not exclusive.

In order to allow for the correct allocation of active substances to the measured residues (see Section 2.4.1.1), this information was integrated in the list of residue definitions. Table 3 provides an overview of all relevant variables.

Table 3: Description of the variables contained in the list of residue definitions

Name	Label	Description
paramCode_RD	Residue code	Code of the residue definition as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
paramName_RD	Residue name	Name of the residue definition as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
paramCode_AS	Substance code	Code of the associated active substance(s) as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
paramName_AS	Substance name	Name of the associated active substance(s) as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
MW_factor	MW conversion factor	Multiplication factor used to convert the amount of measured residue into active substance. This factor is derived from the molecular weights (MW) of both compounds

Name	Label	Description
Is_exclusive	Exclusive	Indicates whether the association between active substance and residue definition is exclusive
Proportion	Proportion	Estimated proportion of the active substance within the associated residue definition, only applicable when the association is not exclusive

RD: residue definition; AS: active substance.

2.2.4. Occurrence data

The occurrence data collected under Article 31 of Regulation (EC) No 396/2005 are the most appropriate data available to EFSA for performing a retrospective exposure assessment to pesticide residues. These data are obtained from the official control activities carried out in the EU Member States, Iceland and Norway. These data are reported to EFSA using the Standard Sample Description (SSD) (EFSA, 2010, 2013). Although the occurrence data are collected at the level of individual measurements, the SSD allows identification of measurements associated to a single food sample (e.g. samples analysed for multiple pesticide residues). After validation by EFSA, the collected data are integrated in the EFSA Data Warehouse.

All occurrence data referring to the relevant food commodities (see Section 2.2.1) and residue definitions (see Section 2.2.3) were extracted from the Data Warehouse. Only measurements validated under the 2014, 2015 and 2016 EU reports on pesticide residues in food were included (EFSA, 2016, 2017, 2018a).

As required by the risk management principles agreed among Member States,[1] the following additional criteria were applied to the extracted data.

- Only samples resulting from the EU-coordinated control programme (EUCP), national control programmes or a combination of those were selected (SSD codes K005A, K009A and K018A). Samples associated to increased control programmes or any other type of program were excluded as they were not considered to be representative.
- Only samples obtained through selective or objective sampling were retained (SSD codes ST10A and ST20A). Samples obtained through suspect sampling or any other type of sampling were not considered to be representative and therefore excluded.
- As the food consumption data are reported for RPCs, samples for processed commodities were excluded from the assessment, except for foods for infants and young children. This means that for the 30 RPCs, only samples with a product treatment specified as 'unprocessed' or 'freezing' were selected (SSD codes T998A and T999A). Regarding foods for infants and young children, the product treatment 'processed' was considered implicit (SSD code T100A).
- Only measurements reported as a numerical (i.e. quantifiable) value or as a non-quantified value were considered useful for the assessment (SSD codes VAL and limit of quantification (LOQ)). Other result types were not considered valid and therefore excluded.
- Only measurements reported for the full legal residue definition or for the most complete subset of the residue definition were used (SSD codes P004A and P005A). Measurements referring to a part of the residue definition were excluded from the assessment.
- When the LOQ value for a measurement could not be reported by the Member States (i.e. for residue definitions composed of multiple components), the median LOQ of all measurements referring to the same combination of commodity and residue definition was assumed.
- When several measurements with overlapping residue definitions were reported for the same sample, only the measurement referring to the most recent residue definition was retained for assessment.

Occurrence data from all EU Member States, Iceland and Norway were pooled into one single data set for each CAG. The key variables retained in the occurrence data set are summarised in Table 4.

Considering the size of the occurrence data sets, only the summary statistics per residue definition and food commodity are reported (see Annex A.1, Table A.1.09 and Annex A.2, Table A.2.09). Occurrence data for water were not available to EFSA and were therefore imputed according to the assumptions elaborated in Section 2.4.1.4.

Table 4: Description of the variables contained in the occurrence data set

Name	Label	Description
labSampCode	Sample code	Alphanumeric code of the analysed sample
prodCode	RPC code	Code of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodName	RPC name	Name of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
paramCode	Residue code	Code of the residue definition as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
paramName	Residue name	Name of the residue definition as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
resLOQ	Limit of quantification	The lowest quantifiable amount (in mg/kg) detectable by the laboratory's analytical system
resVal	Result value	Concentration of the measured residue (in mg/kg) within the analysed sample
resType	Result type	Indicates the type of result, whether it could be quantified/determined or not

RPC: raw primary commodity.

2.2.5. Consumption data

The EFSA Comprehensive European Food Consumption Database (Comprehensive Database) provides a compilation of existing national information on food consumption at individual level. It was first built in 2010 (EFSA, 2011; Huybrechts et al., 2011; Merten et al., 2011). Details on how the Comprehensive Database is used are published in the Guidance of EFSA (EFSA, 2011). Data reported in the Comprehensive Database may either refer to RPCs, RPC derivatives (i.e. single-component foods altered by processing) or composite foods (i.e. multicomponent). Consumption data for RPC derivatives and composite foods, however, cannot be used in exposure assessments when the occurrence data are reported for the RPCs.

To address the above issue, EFSA transformed the Comprehensive Database into a new RPC Consumption Database by means of the RPC model (EFSA, 2019d). This model converts the consumption data for composite foods or RPC derivatives into their equivalent quantities of RPCs. The RPC model was applied to the Comprehensive Database as of 31 March 2018, when it contained results from 51 different dietary surveys carried out in 23 different Member States covering 94,523 individuals.

In view of the current pilot project, the food consumption data extracted from the RPC Consumption Database were limited to the population classes and countries listed below.

- Toddlers[2]: Denmark, the Netherlands and the United Kingdom
- Other children[3]: Bulgaria, France and the Netherlands
- Adults[4]: Belgium, the Czech Republic, Italy and Germany

An overview of the selected dietary surveys is provided in Annex A.1, Table A.1.04 and Annex A.2, Table A.2.04.

Using the data extraction process, a single consumption data set was obtained for acute exposure assessment. Hence, the same data were used for assessment of both CAG-NAN and CAG-NAM. The key variables retained in the occurrence data set are summarised in Table 5. Summary statistics on the quantities of RPC consumed per country, survey and population class are reported (see Annex A.1, Table A.1.10 and Annex A.2, Table A.2.10).

---

[2] The population class 'toddlers' refers to participants from $\geq$ 12 months to < 36 months old.
[3] The population class 'other children' refers to participants from $\geq$ 36 months to < 10 years old.
[4] The population class 'adults' refers to participants from $\geq$ 18 years to < 65 years old.

Table 5: Description of the variables contained in the food consumption data set

Name	Label	Description
Country	Country	Country where the dietary survey took place as defined by EFSA's harmonised terminology for scientific research (COUNTRY catalogue; EFSA, 2019c)
Survey	Survey	Acronym of the dietary survey
PopClass	Population class	Participant's population class, based on age, as defined by EFSA's harmonised terminology for scientific research (AGECLS catalogue; EFSA, 2019c)
ORSUBID	Subject ID	A pseudonymised subject ID number generated by EFSA upon receipt of the data
Weight	Body weight	Bodyweight of the subject (in kg)
ndays	Number of survey days	Number of days on which the participant's consumption was surveyed
day	Survey day	Ordinal number of the day on which the participant's consumption was surveyed
prodCode	RPC code	Code of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodName	RPC name	Name of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
FoodEx2_Facets	Processing code	FoodEx2 facet code describing the processing technique, including additional descriptors such as qualitative information, part consumed or the nature of the food (EFSA, 2015b)
RPCD_amount	RPCD amount	Amount of raw primary commodity derivative (in grams)
RPC_amount	RPC amount	Amount of raw primary commodity (in grams)

RPC: raw primary commodity; RPCD: raw primary commodity derivative.

2.3. Secondary input data

2.3.1. Maximum residue levels

Certain assumptions on the extrapolation of occurrence data (see Section 2.4.1.2) require information on the maximum residue levels (MRLs). An MRL is the upper legal level of a concentration for a pesticide residue in or on food or feed set in accordance with Regulation (EC) No 396/2005. This regulation also defines a procedure for the setting and modification of MRLs. MRLs may therefore have been modified throughout the 2014–2016 reference period. In order to obtain a single list of MRLs, EFSA decided to use the MRLs as of 31 December 2016 (i.e. the end of the current reference period). Hence, it was assumed that those MRLs were applicable during the entire reference period, regardless whether the MRL or residue definition may have changed during that period.

MRLs for the relevant food commodities (see Section 2.2.1) and residue definitions (see Section 2.2.3) were extracted from the EU Pesticides Database[5] and organised in a data format that can be used directly for exposure assessment (see Annex A.1, Table A.1.05 and Annex A.2, Table A.2.05). Table 6 describes the variables that were part of this data format.

Table 6: Description of the variables contained in the list of maximum residue levels

Name	Label	Description
paramCode_RD	Residue code	Code of the residue definition as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
paramName_RD	Residue name	Name of the residue definition as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
prodCode	RPC code	Code of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)

---

[5] http://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/public/?event=homepage&language=EN

Name	Label	Description
prodName	RPC name	Name of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
MRL	MRL (mg/kg)	Numerical value of the maximum residue level (MRL) as defined by Regulation (EC) No 396/2005, expressed in mg/kg
atLOQ	MRL at LOQ	Indicates whether the maximum residue level (MRL) is set at the analytical limit of quantification (LOQ). Under Regulation (EC) No 396/2005 such MRLs are marked with an asterisk (*)

RD: residue definition; RPC: raw primary commodity.

2.3.2. Authorised uses

In some cases, the imputations and simulations performed on the occurrence data rely on the authorisations for use of the active substance(s) (see Section 2.4.1). While the approval status of an active substance under Regulation (EC) No 1107/2009 is regulated at EU level, the authorisations for use of active substances on specific crops are delivered at national level within the EU Member States. A centralised database compiling these national authorisations is not yet available at EU level.

National authorisations can be reported to EFSA under Regulation (EC) No 396/2005, either for an MRL application under Article 10, or for an MRL review under Article 12. There is, however, no legal obligation to systematically report all national authorisations and the MRL review programme is still in progress. A comprehensive overview of all pesticide authorisation within the EU is therefore also not available to EFSA. Meanwhile, a tentative list of authorised uses was elaborated according to the following principles.

- When the MRL for a given combination of active substance and RPC was not set at the LOQ (see Section 2.3.1), the active substance was assumed to be authorised for use on that specific commodity. This assumption also accounts for uses authorised outside the EU and for which treated products may be placed on the EU market.
- For the group of dithiocarbamates, which comprises six active substances, Regulation (EC) No 396/2005 provides specific information on the active substances that were used for deriving the MRLs. Authorised uses for these active substances were identified accordingly.
- For the remaining combinations of active substance and RPC (i.e. where the MRL was set at LOQ), EFSA screened the relevant reasoned opinions issued under Article 12 of Regulation (EC) No 396/2005 and the subsequent reasoned opinions issued under Article 10. Any authorised use reported in those reasoned opinions was recorded.
- When the MRL was set at LOQ and a review under Article 12 of Regulation (EC) No 396/2005 had not been issued, it was assumed that the use was not authorised.

The authorised uses collected by EFSA were integrated in a data format that can be readily used for exposure assessment (see Annex A.1, Table A.1.06 and Annex A.2, Table A.2.06). Table 7 describes the variables of this data format.

Table 7: Description of the variables contained in the list of authorised uses

Name	Label	Description
paramCode_AS	Substance code	Code of the active substance as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
paramName_AS	Substance name	Name of the active substance as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
prodCode	RPC code	Code of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodName	RPC name	Name of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
Source	Source	Indicates the source of the information (i.e. MRL legislation, MRL review or MRL application)
Reference	Reference	EFSA Journal reference to the relevant reasoned opinion (i.e. when the information was retrieved from an MRL review or application)

AS: active substance; RPC: raw primary commodity; MRL: maximum residue level.

2.3.3. Extrapolation rules

The extrapolation of occurrence data described in Section 2.4.1.2 is carried out in compliance with the guidelines on comparability, extrapolation, group tolerances and data requirements for setting MRLs (European Commission, 2017). These extrapolation rules depend on when the active substance is applied to the plant.

For the current assessment, available occurrence data do not provide any information on how the plant commodity was treated. Therefore, the most conservative extrapolation rules were applied, i.e. for treatments after formation of the edible plant parts. These extrapolation rules were integrated in a data format that can be readily used for exposure assessment (see Annex A.1, Table A.1.07 and Annex A.2, Table A.2.07). Table 8 describes the variables of this data format.

Table 8: Description of the variables contained in the list of extrapolation rules

Name	Label	Description
prodCode_from	RPC code (from)	Code of the raw primary commodity from which the extrapolated measurements are taken (i.e. source commodity). This code is compliant with EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodName_from	RPC name (from)	Name of the raw primary commodity from which the extrapolated measurements are taken (i.e. source commodity). This code is compliant with EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodCode_to	RPC code (to)	Code of the raw primary commodity to which the measurements are extrapolated (i.e. target commodity). This code is compliant with EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodName_to	RPC name (to)	Name of the raw primary commodity to which the measurements are extrapolated (i.e. target commodity). This code is compliant with EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)

RPC: raw primary commodity.

2.3.4. Processing factors

Occurrence data for pesticide residues are collected at the level of RPC (see Section 2.2.4). Food consumption data may be collected at the level of RPC, RPC derivative or composite food, but for the purpose of this assessment all consumption data for composite foods and RPC derivatives were converted into their equivalent quantities of RPCs (see Section 2.2.5). Combining occurrence and consumption data at the RPC level implies that all residues present in the RPC will reach the end consumer. This assumption, however, is very conservative. In reality, these residues will most likely be altered through processing, such as peeling, washing, cooking, etc.

The effect of processing is usually addressed by means of processing factors. A processing factor is specific to each RPC, processing type and active substance, and it accounts for both the chemical alteration of the substance and weight change of the food. Processing factors are quantified by dividing the expected residue concentration in the processed commodity by the residue concentration in the raw commodity.

The European database on processing factors is the most recent and the most comprehensive compilation of processing factors currently available at EU level (Scholz et al., 2018). Processing factors for the active substances and RPCs under assessment were extracted from the database according to the following criteria.

- For each active substance, RPC and processing technique only the median processing factor was extracted.
- Only the processing factors indicated as reliable or indicative were extracted. Processing factors indicated as unreliable were excluded from the assessment.

Processing techniques reported in the processing factor database were then compared to the processing techniques reported in the RPC consumption data set. The processing techniques from both databases were matched according to the following principles:

- When a generic processing technique was reported in the RPC consumption database (e.g. juice) while more specific processing techniques were reported in the processing factor database (e.g. pasteurised juice and unpasteurised juice), the specific processing technique with the highest processing factor was selected.
- When a specific processing technique was reported in the RPC consumption database (e.g. mashed potato) while a more generic processing technique was reported in the processing factor database (e.g. boiled potato), the generic processing factor was applied to the specific processing techniques.
- When a processing factor was reported for an isomeric mixture (e.g. benalaxyl), the processing factor was considered valid for any isomeric mixture that contained the same isomers (e.g. benalaxyl-M).
- Processing factors were extrapolated between RPCs with similar properties (i.e. oranges and mandarins, apples and pears, table and wine grapes, wheat and rye grain).
- Processing factors for peeling were applied to the corresponding fruit with inedible peel, even when the processing technique was not specified in the RPC consumption database (i.e. oranges, mandarins, bananas and melons).

By following these principles, lists of processing factors were obtained for the assessment of both CAG-NAN and CAG-NAM (see Annex A.1, Table A.1.08 and Annex A.2, Table A.2.08, respectively).

Table 9 describes the variables contained in the list of processing factors.

Table 9: Description of the variables contained in the list of processing factors

Name	Label	Description
paramCode_AS	Substance code	Code of the active substance as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
paramName_AS	Substance name	Name of the active substance as defined by EFSA's harmonised terminology for scientific research (PARAM catalogue; EFSA, 2019c)
prodCode	RPC code	Code of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodName	RPC name	Name of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
facetCode	Processing code	FoodEx2 facet code describing the processing technique, including additional descriptors such as qualitative information, part consumed or the nature of the food (EFSA, 2015b)
facetDesc	Processing description	Description of the processing code
procFac	Processing factor	Numerical value representing the expected residue concentration in the processed commodity divided by the residue concentration in the raw commodity
Source	Source	Indicates the source of the information (i.e. type of report)
Reference	Reference	Journal reference to the relevant report
Comment_PF	Comment	Indicates whether the processing factor relies on any type of assumption or extrapolation

RPC: raw primary commodity; PF: processing factor.

2.3.5. Variability factors

Acute exposure assessments for pesticide residues should account for variability among the single commodity units of the composite laboratory samples (see Section 2.4.2). To account for this variability, several parameters are required for each food commodity.

- Unit weight: estimated weight for a single commodity unit.
- Units per sample: estimated number of units within a composite laboratory sample.
- Variability factor (VF): expected variability among the single unit concentrations, which is defined as the ratio between the 97.5th percentile and mean of the distribution of unit concentrations.

Unit weights for each commodity were retrieved from the Pesticide Residues Intake Model (EFSA, 2018b). Commodities with a unit weight inferior to 25 g were not further considered because, in this

case, the residue concentration in the composite laboratory sample is expected to reflect the residue concentration in the portion that would be consumed (FAO, 2003).

The number of units per sample was obtained from Commission Directive 2002/63/EEC, establishing community methods of sampling for the official control of pesticide residues in and on products of plant and animal origin. This directive defines a minimum weight and a minimum number of units for composite laboratory samples of each food category. Hence, the minimum number of units (as defined by Directive 2002/63/EEC) was used, unless the minimum sample weight divided by the corresponding unit weight was higher. In that case, the latter calculated value (rounded up to the next integer) was retained.

VFs were also retrieved from the Pesticide Residues Intake Model (EFSA, 2018b). According to the risk management principles agreed among Member States,[1] these factors were only used for the Tier I scenario. For the Tier II scenario, a VF of 3.6 was applied to all commodities having a unit weight above 25 g (see also Section 2.4.2).

As the above-mentioned parameters are defined by food commodity, the relevant variables were incorporated in the list of food commodities (see Annex A.1, Table A.1.02 and Annex A.2, Table A.2.02). Table 10 provides a complete overview of all the variables contained in the list of food commodities.

Table 10: Description of the variables contained in the list of raw primary commodities

Name	Label	Description
prodCode	RPC code	Code of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
prodName	RPC name	Name of the raw primary commodity as defined by EFSA's harmonised terminology for scientific research (MATRIX catalogue; EFSA, 2019c)
Cat_2002_63_EC	Category Directive 2002/63/EC	Commodity classification defined by Table 4 of the Annex to Commission Directive 2002/63/EC
SampWeight	Minimum sample weight	Minimum size of each laboratory sample (expressed in kg) defined by Table 4 of the Annex to Commission Directive 2002/63/EC
minUnits	Minimum units per sample	Minimum size of each laboratory sample (expressed in number of units) defined by Table 4 of the Annex to Commission Directive 2002/63/EC
UnitWeight	Unit weight	Estimated weight (expressed in g) for a single commodity unit as reported in the Pesticide Residues Intake Model (EFSA, 2018b)
NrUnits	Units per sample	Estimated number of units required to obtain the minimum size of a laboratory sample, both in terms of weight and number of units
VF	Variability factor	Default variability factor as reported in the Pesticide Residues Intake Model (EFSA, 2018b). This factor represents the variability among the single unit concentrations, which is defined as the ratio between the 97.5th percentile and mean of the distribution of unit concentrations

RPC: raw primary commodity.

2.3.6. Processing types

Variability among the single commodity units of the composite laboratory samples is not relevant when the food consumed is subject to processing techniques that involve bulking and blending.

EFSA therefore extracted all processing techniques reported in the RPC consumption data (see Section 2.2.5) and identified the processes that normally involve blending or bulking. Typically, these are processing techniques performed at industrial level (e.g. milling, oil production, etc.). Household processes, however, were assumed not to involve any bulking or blending (e.g. boiling, stewing, etc.). However, although juicing may also be carried out at household level, EFSA assumed that most fruit juices are produced at industrial level.

The same list of processing types was used for the assessment of both CAGs (see Annex A.1, Table A.1.09 and Annex A.2, Table A.2.09). Table 11 describes the variables contained in the list of processing types.

Table 11: Description of the variables contained in the list of processing types

Name	Label	Description
Facets	Processing code	FoodEx2 facet code describing the processing technique, including additional descriptors such as qualitative information, part consumed or the nature of the food (EFSA, 2015b)
Facets_desc	Processing description	Description of the processing code
Blending	Blending or bulking	Indicates whether the processing technique involves any type of blending or bulking

2.4. Inner-loop execution

2.4.1. Adjustments and simulations on the occurrence data

2.4.1.1. Allocation of active substances to the occurrence data

While the CAGs are defined at the level of the pesticide active substances, the occurrence data reported to EFSA refer to residue definitions for enforcement purposes (see Section 2.2.4). Hence, the original occurrence data set obtained from the EFSA Data Warehouse is converted into a new intermediate data set where measurements are assigned to active substances instead of residue definitions.

Some of these residue definitions, however, referred to as unspecific residue definitions, may be associated to multiple active substances (see Section 2.2.3). Allocation of active substances to these unspecific residue definitions is performed in compliance with the risk management principles agreed among Member States.[1]

Under the Tier I assumptions, measurements for unspecific residue definitions are always assigned to the most potent active substance (i.e. the substance with the lowest NOAEL), regardless of its authorisation status. This approach is expected to overestimate the exposure because a less potent active substance may have been used. This overestimation may be even more substantial when the most potent active substance is not authorised for use on the relevant commodity.

A more likely scenario would be the use of a combination of more potent and less potent substances. Therefore, for the Tier II calculations, each measurement is randomly assigned to one of the active substances authorised on that commodity, regardless of whether the active substance is part of the CAG or not. If none of the active substances associated to the unspecific residue definition is authorised, any active substance is selected at random. Furthermore, special consideration is given to the active substances that may metabolise into another active substance, the non-exclusive substances (see Section 2.2.3). If the measurement is assigned to a non-exclusive substance (e.g. dimethoate), the model assumes that the measurement is partially composed of the assigned active substance while the remaining fraction is attributed to the active substance into which it metabolises (e.g. omethoate), the exclusive substance.

A more detailed description of the methodologies used to allocate active substances to the occurrence data is provided in Appendix A.

Although the Tier II assumptions are expected to better reflect reality, some uncertainties related to this approach were still identified. Under ideal circumstances, the probability to select an active substance should be based on market share data for those active substances. Similarly, the proportion of the non-exclusive substance should be derived from the available metabolism data. Both market share data and metabolism data, however, were not readily available. In the absence of these data, assumptions on equal probability and equal proportion are applied instead. It should be noted that these assumptions may either underestimate or overestimate the exposure.

An additional uncertainty derives from the assumption that measurements for unspecific residue definitions result from the use of single active substances. This assumption implies that other active substances associated to that unspecific residue definition are not present (i.e. implicit zero measurements). Although it is unlikely that substances with similar pesticidal activity are used on the same crop, this possibility cannot be excluded.

2.4.1.2. Extrapolation of occurrence data

For some active substances and food commodities, the number of measurements may be limited. Furthermore, for certain combinations, data may even be missing completely. In order to address the

uncertainties related to those limited or missing data, extrapolation rules are integrated in the exposure model.

The extrapolations are carried out in compliance with the guidelines on comparability, extrapolation, group tolerances and data requirements for setting MRLs (European Commission, 2017). These extrapolation rules were developed and validated in view of extrapolating occurrence data from 'data rich' commodities (e.g. apples) to 'data poor' commodities (e.g. pears). However, there is currently no clear guidance on the number of measurements needed to perform a reliable probabilistic calculation. In the framework of this assessment, a minimum of 10 measurements per active substance and commodity is considered sufficient to perform a reliable probabilistic assessment.

Hence, only the combinations of food commodities and active substances with less than 10 measurements are extracted from the data set (i.e. the target combination). Measurements for the same active substance but a different commodity (i.e. source commodity) are then extrapolated to the target commodity provided that:

- the extrapolation is compliant with the extrapolation rules reported in Section 2.3.3;
- the MRLs are the same in both the source and target commodity;
- the use of the active substance is authorised in both the source and target commodity; and
- the number of measurements for the source commodity is higher than or equal to 10.

The extrapolated measurements are randomly assigned to the available target commodity samples, excluding samples where the active substance was already measured. The number of extrapolated measurements is reported in the final output (see Section 2.6).

The methodology used for extrapolation of occurrence data is independent of the Tier I or Tier II scenarios. A more detailed description is provided in Appendix B.

2.4.1.3. Imputation of left-censored occurrence data

Over 95% of the occurrence data used for the current exposure assessment are left-censored (see Section 2.2.4). Left-censored data are measurements reported below the LOQ and for which an accurate value is not available. Some of these results may be low-positive residues while others will be true zeroes (no-residue situation).

In order to address the uncertainties resulting from the high proportion of left-censored data, measurements below the LOQ were imputed in compliance with the risk management principles agreed among Member States.[1]

Under Tier I assumptions, left-censored measurements were imputed with 1/2 LOQ when at least one positive result (i.e. above LOQ) was reported for a given substance-commodity combination. Measurements for all remaining combinations were imputed with a zero (i.e. assuming a no-residue situation).

For the Tier II assessment, use frequencies are estimated for each pesticide and each commodity, assuming that all samples were treated according to at least one agricultural use pattern (AUP).[6] An AUP is the combination of pesticide uses applied to a single commodity or crop. The estimated use frequencies are then used to calculate a proportion of true zeros and the corresponding number of left-censored measurements is then selected at random from the data set. While the selected measurements are imputed with zero, the remaining left-censored measurements are imputed with 1/2 LOQ. A more detailed description of the methodology is provided in Appendix C.

As for the allocation of active substances (see Section 2.4.1.1), the Tier II assumption is expected to more refined compared to the Tier I assumption, which is a very conservative assumption. These Tier II calculations would be even more accurate if actual data on the use frequency of pesticides would be made available. In particular, for pesticides with unquantifiable residues the estimated use frequency will be 0% which is most likely an underestimation of the real use frequency. On the other hand, this scenario also assumes that the total AUP frequency is 100%, meaning that all commodities were treated according to at least one AUP. This tends to overestimate the exposure.

2.4.1.4. Imputation of missing occurrence data

In acute cumulative exposure assessments, it is necessary to take account of any correlations that may exist between the concentrations of the different active substances within a given food sample (EFSA PPR Panel, 2012). Under the current assessment, the correlations of chemicals within a single

---

[6] The method described is an interpretation of 'Option 5' defined by the risk management principles on cumulative risk assessment (http://www.efsa.europa.eu/sites/default/files/SANTE_CRA_Mandate.pdf).

sample are accounted for because measurements obtained from the same sample are identified in the available occurrence data (see Section 2.2.4). The samples, however, were not necessarily analysed for every substance of the CAG. Measurements within a sample may therefore be missing for some substances. The number of missing values for each substance-commodity combination can be retrieved from Annex A.1, Table A.1.10 and Annex A.2, Table A.2.10.

In order to avoid underestimation of the acute cumulative exposure, missing measurements are imputed according to the risk management principles agreed among Member States.[1]

For each substance-commodity combination, the number of missing values is counted, and the same number of measurements is randomly selected from the available data set. The missing values are then replaced with the selected measurements. The Tier I scenario uses a very conservative assumption where the missing value of the most contaminated sample is imputed with the highest possible imputation value. Under the Tier II scenario, a more refined assumption is applied where imputation values are assigned at random. A more detailed description of the methodology is provided in Appendix D.

2.4.1.5. Imputation of occurrence data for water

Occurrence data for water are not available to EFSA (see Section 2.2.4). As required by the risk management principles agreed among Member States,[1] occurrence data for water are imputed for the five most potent active substances within the CAG.

For this purpose, the five substances with the lowest NOAEL are extracted from the list of active substances (see Section 2.2.1) and a measurement in water is added to the occurrence data set for each of these substances. These measurements are associated to a single fictitious sample code. While under the Tier I assessment, a result value of 0.001 mg/kg is assigned to each measurement, a result value of 0.0005 mg/kg is assigned under Tier II.

2.4.2. Acute exposure distribution

Acute dietary exposure is modelled at the level of individual consumption days by means of an empirical Monte Carlo simulation (EFSA PPR Panel, 2012). This means that individual days are selected at random from the consumption data set. For each food commodity consumed within the individual days, random samples of the occurrence data set are assigned. Using the individuals' bodyweights and the concentration of the different active substances measured in the different samples, the acute exposures resulting from each food commodity and active substance within each individual day are calculated.

The occurrence data used for the assessment, however, relate to the average concentrations in composite laboratory samples (see Section 2.2.4). Consumers on the other hand are exposed to individual units of the commodity. Residue concentrations may vary among the individual units, referred to as unit-to-unit variability. For RPCs that have a unit weight inferior to 25 g and for processed foods that were subject to blending or bulking, the unit-to-unit variability is not considered relevant (FAO, 2003). For the remaining food commodities, a fixed 'variability factor' is usually applied for acute deterministic calculations. The VF is defined as the ratio between the 97.5th percentile and mean of the distribution of unit concentrations.

For probabilistic exposure assessment, the use of a distribution of unit concentrations is considered more adequate than using a fixed VF. Therefore unit-to-unit variability is modelled using a beta distribution, which can be bounded between 0 and an upper limit. Indeed, if the average concentration in a composite sample is 1, the concentration in a single unit can never be higher than the number of units within the composite sample (assuming all other units have a concentration of zero). Hence, for each RPC with a unit weight exceeding 25 g, the beta distribution was parametrised with the following restrictions.

- Lower bound = 0
- Mean = 1
- 97.5th percentile = variability factor
- Upper bound = number of units per sample.

Stochastic variability factors (SVF) can then be drawn from the beta distribution and multiplied with the composite sample concentration to obtain a plausible estimate of the unit concentration. When the portion consumed by an individual is smaller than a single unit, the SVF is directly applicable to the consumed portion. When the consumed portion is composed of multiple units, however, multiple SVFs will be drawn from the same beta distribution in order to estimate concentration in the whole portion

consumed. Therefore, the concentration in the whole portion is estimated by multiplying the sample concentration with a weighted variability factor (WVF), which is calculated as follows.

$$WVF = SVF_n \qquad \text{if } n = 1$$

$$WVF = \frac{\left(\sum_{i=1}^{n-1} SVF_i\right) + SVF_n \cdot (n_0 - n + 1)}{n_0} \quad \text{if } n > 1$$

where WVF is the weighted variability factor;
SVF_i is the stochastic variability factor drawn for unit i;
n_0 is the estimated number of units within the consumed portion (unrounded);
n is the number of stochastic variability factors to be drawn (i.e. ceiling of n_0).

In order to combine the different substances in a total acute exposure estimate, the toxicological potency of each substance also needs to be accounted for. The use of relative potency factors has previously been suggested by EFSA (EFSA PPR Panel, 2012) but this method requires identification of an index compound for each CAG. Alternatively, the exposure estimates for the individual active substances are divided by the corresponding NOAEL. The potency-adjusted estimates can then be combined to obtain a total normalised exposure (NET) for each individual day.

Combining occurrence and consumption data at the RPC level also implies that all residues present in the RPC will reach the end consumer, while alteration of residues is expected to occur when the RPCs are processed prior to consumption. This uncertainty, which is generally expected to overestimate exposure, is addressed by integrating processing factors where available (see Section 2.3.4). Considering, however, that processing factors account for both the chemical alteration of the substance and weight change of the food, occurrence values need to be combined with the consumed amount of processed food (i.e. RPC derivative) instead of the consumed amount of RPC. Furthermore, as the consumed amounts are expressed in g and occurrence data are expressed in mg/kg, a correction factor of 1,000 needs to be considered.

Based on the considerations above, the NET is calculated for each individual day according to the equations reported below.

$$NET_{id} = \sum_{c}^{\text{Commodities}} \sum_{p}^{\text{Processes}} \sum_{s}^{\text{Substances}} \begin{cases} \dfrac{RPC_{idcp} \cdot X_{idcps} \cdot WVF_{idcps}}{BW_i \cdot NOAEL_s \cdot 10^3} & \text{if } PF_{cps} \text{ unspecified} \\[2ex] \dfrac{RPCD_{idcp} \cdot X_{idcps} \cdot WVF_{idcps} \cdot PF_{cps}}{BW_i \cdot NOAEL_s \cdot 10^3} & \text{if } PF_{cps} \text{ specified} \end{cases}$$

where NET_{id} is the total normalised exposure of individual i on day d;
RPC_{idcp} is the amount of commodity c with processing type *p* consumed by individual i on day d, expressed in g of raw primary commodity;
$RPCD_{idcp}$ is the amount of commodity c with processing type p consumed by individual i on day d, expressed in g of raw primary commodity derivative;
BW_i is the body weight of individual i, expressed in kg;
X_{idcps} is the average concentration of substance s in the sample that was randomly assigned to individual i on day d for commodity c with processing type p, expressed in mg/kg;
WVF_{idcps} is the weighted variability factor that was randomly assigned to individual i on day d for substance s in commodity c with processing type p;
PF_{cps} is the processing factor for substance s in commodity c with processing type p;
$NOAEL_s$ is the no observed adverse effect level for substance s, expressed in mg/kg bodyweight;

The Monte Carlo simulation described above is performed with 100,000 iterations. This means that, for each dietary survey, 100,000 individual days are randomly selected with replacement and NETs are calculated for each individual day. This results in empirical distributions of NETs, representing the variability of single day exposures within the different population groups.

The methodology used to derive the acute exposure distribution is generally the same for both scenarios (Tier I and Tier II). The only difference lies in the beta-distribution used to reflect unit-to-unit variability. While a very conservative VF of 5 or 7 is used for the Tier I scenario (see also

Section 2.2.5), a fixed VF of 3.6 is applied under the Tier II scenario. A more detailed description of the methodology used to estimate acute dietary exposure is provided in Appendix E.

2.5. Outer-loop execution

The consumption data used for this assessment are subject to sampling uncertainty and will not represent perfectly the true diets within the population. Likewise, the occurrence data will not perfectly reflect the true distribution of residue concentrations in food. These sampling uncertainties are addressed by repeating the inner-loop execution multiple times, each time replacing the consumption and occurrence data sets with bootstrap data sets (EFSA PPR Panel, 2012). Bootstrap data sets are obtained by resampling, with replacement, the same number of observations from the original data sets. Each time the inner-loop is executed with bootstrap data sets, a bootstrap distribution of NETs will be obtained. This shows how the distribution of NETs may have looked like if random sampling from the population would have generated different samples than those actually observed (Efron and Tibshirani, 1993).

It should be noted, however, that the both the consumption and occurrence data incorporate several multivariate patterns (e.g. association of foods and individuals' characteristics, co-occurrence of residues, etc.). These patterns need to be preserved in the bootstrap data sets.

Consumption data are, therefore, resampled at the individual day level, i.e. selecting all consumption events of the resampled individual day. Hence, for each dietary survey, the bootstrap data sets contain the same number of individual days as the original data set.

Occurrence data, on the other hand, are resampled at the level of the laboratory sample, i.e. selecting all measurements obtained in the resampled laboratory sample. Hence, the bootstrap data sets contain for each food commodity the same number of laboratory samples as the original data set.

In the current exposure model, the inner-loop execution is repeated 100 times. The first execution, also referred to as the nominal run, is performed with the original data sets. The remaining executions are performed with bootstrap data sets.

Although the outer-loop execution is primarily intended to address the sampling uncertainty of the consumption and occurrence data, it also addresses uncertainty resulting from the probabilities applied in the model. This is particularly true for the Tier II scenarios where several simulations and imputations rely on the random selection of measurements (see Section 2.4.1).

2.6. Output preparation

Through the inner- and outer-loop executions, multiple NET distributions are generated (i.e. 100 bootstrap distributions per dietary survey). To describe each bootstrap distribution, the following parameters are derived:

- mean of the NET;
- standard deviation of the NET;
- percentiles of the NET (P2.5, P5, P10, P25, P50, P75, P90, P95, P97.5, P99, P99.9 and P99.99).

As required by the risk management principles agreed among Member States,[1] the parameters of the exposure distribution are expressed in total margin of exposure (MOET). The margin of exposure is normally calculated as the ratio of a toxicological reference dose (i.e. NOAEL) to the estimated exposure. Considering that the exposure is already normalised (see Section 2.4.2), the MOET is in this case the reciprocal value of the NET.

As a result, 100 MOET estimates are obtained for each parameter of the exposure distributions. These 100 estimates reflect the uncertainty distribution around the true value of those parameters. From these uncertainty distributions a 95% confidence interval is calculated for each parameter. The median of the uncertainty distribution is selected as the central estimate for the confidence interval.

To better understand the factors that influence the lowest MOETs (or the highest NETs), individual days with an MOET lower than the MOET calculated at the 99th percentile of the exposure distribution are extracted for the nominal run of each dietary survey. The relevant information associated to those individual days is also retrieved (i.e. amounts of foods consumed and concentrations of active substances). Based on the individual days' information, average contributions are calculated per dietary survey, active substance and food commodity.

Additional information is gathered throughout the calculation process to support the identification of missing information. These intermediate outputs mainly refer to the missing occurrence data and possible extrapolations (see Section 2.4.1.2). For the Tier II scenario, the estimated use frequencies are also reported (see Section 2.4.1.3).

The above reported percentiles were calculated using SAS® software, which provides five validated options for the definition of percentiles.[7] For the purpose of this assessment, the following percentile definition was selected. Let n be the number of non-missing values for a variable, let $x_1, x_2, \ldots, x_n$ represent the ordered values of the variable and set $p = t/100$. Then the tth percentile is calculated as follows.

$$y = (1 - g)x_j + gx_{j+1}$$

where y   is the tth percentile;

       j   is the integer part of np;

       g   is the fractional part of np.

The percentile definition is not expected to have a substantial impact for the acute exposure estimates because 100,000 individual days are simulated for each exposure distribution. With such a high number of observations, calculated percentiles are expected to be stable regardless of the percentile definition used.

2.7. Tiers and sensitivity analyses

As required by the risk management principles agreed among Member States,[1] the methodology described above is applied in a tiered approach:

1) The Tier I scenario uses very conservative assumptions that are less resourceful regarding data and computational capacity. This allows for an efficient screening of the exposure with low risk for underestimation of the real exposure to pesticide residues.

2) The Tier II scenario, which is more resourceful, includes more refined assumptions but it is still intended to be conservative.

Table 12 summarises the main assumptions and methodologies applied in the exposure model. The key differences between Tier I and Tier II are also highlighted. Although the methods and assumptions applied in the model were selected with the view of minimising the uncertainties, resources may sometimes be insufficient to allow for a more accurate assessment (e.g. use frequencies and processing factors). In order to assess how these additional data or improvement might impact on the exposure estimates, the following sensitivity analyses were also carried out:

- Sensitivity analysis A assumes that left-censored data are imputed at 1/2 LOQ when the use of the active substance is authorised.
- Sensitivity analysis B assumes that all left-censored data are imputed at zero.
- Sensitivity analysis C assumes that residues will not be present in any processed food.
- Sensitivity analysis D excludes all foods for infants and young children.

For these sensitivity analyses, only the impact on the 99.9th percentile of the exposure distribution (expressed in MOET) was assessed. Detailed results were in this case not provided.

---

[7] http://support.sas.com/documentation/cdl/en/procstat/66703/HTML/default/viewer.htm#procstat_univariate_details13.htm

Table 12: Overview of the main assumptions and methodological approaches used for assessing acute cumulative exposure to pesticide residues

	Description	
Consumption data		
Number of surveys	10	
Population classes	Adults (Belgium, Czech Republic, Germany and Italy) Other children (Bulgaria, France and Netherlands) Toddlers (Denmark, Netherlands and United Kingdom)	
Food commodities	30 raw primary commodities (includes conversion from foods as eaten) + 4 categories of foods for infants and young children + water	
Occurrence data (extraction)		
Reference period	2014–2016 (latest available 3-year cycle)	
Food commodities	30 raw primary commodities (unprocessed or frozen) + 4 categories of foods for infants and young children	
Residue definitions	All residue definitions associated to CAG-NAN and CAG-NAM during the reference period (excl. overlapping residue definitions at sample level)	
Sampling framework	EU-coordinated or national control programmes	
Sampling type	Objective or selective sampling only	
Occurrence data (simulations and imputations)		
Unspecific residue definitions	Tier I: Most potent active substance is allocated to each sample	Tier II: Random allocation of authorised active substances to each sample*
Extrapolations	Extrapolation of measurements per active substance and commodity in accordance with guidance document SANCO 7525/VI/95 (European Commission, 2017), when MRL is equal and substance is authorised in both source (N ≥ 10) and target (N < 10) commodities	
Left-censored data	Tier I: Imputed at 1/2 LOQ for food-substance combinations with quantifiable findings	Tier II: Imputed at 1/2 LOQ based on estimated use frequencies (assuming 100% crop treatment)
Missing measurements	Tier I: Highest values assigned to the most contaminated samples	Tier II: Random assignment of missing measurements to available samples
Drinking water	Tier I: Imputed at 0.1 μg/l for the 5 most potent active substances	Tier II: Imputed at 0.05 μg/L for the 5 most potent active substances
Exposure calculations		
Exposure model	Empirical Monte Carlo simulation (inner-loop execution, n = 100,000)	
Uncertainty model	Empirical bootstrapping (outer-loop execution, n = 100)	
Processed foods	Processing factors obtained or extrapolated from the European database on processing factors for pesticides in food (Scholz et al., 2018)	
Unit-to-unit variability	Tier I: Unit concentration sampled from beta-distribution with variability factors defined by the Pesticide Residues Intake Model (PRIMo)	Tier II: Unit concentration sampled from beta-distribution with a variability factor of 3.6

CAG-NAN: cumulative assessment group of pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition; CAG-NAM: cumulative assessment group of pesticides associated with functional alterations of the motor division; MRL: maximum residue level; LOQ: limit of quantification.

*: Accounts for substances that are not part of the CAG and for residue definitions that are not exclusive (see Section 2.4.1.1).

3. Results

The results section summarises the acute cumulative exposure estimates obtained from the calculations. Exposure estimates are presented for 2 CAGs (CAG-NAN and CAG-NAM), 2 different scenarios (Tier I and Tier II) and 10 different dietary surveys. More detailed results (including graphs and charts) are provided in the annexes.

- Annex B.1 presents the results of the Tier I cumulative exposure calculations to CAG-NAN.
- Annex B.2 presents the results of the Tier I cumulative exposure calculations to CAG-NAM.
- Annex C.1 presents the results of the Tier II cumulative exposure calculations to CAG-NAN.
- Annex C.2 presents the results of the Tier II cumulative exposure calculations to CAG-NAM.

All exposure estimates are expressed in MOET, which is the ratio of a toxicological reference dose (i.e. NOAEL) to the estimated exposure (see Section 2.4.2). Hence, an MOET below 1 implies that the estimated exposure exceeds the NOAEL. Likewise, an MOET of 100 means that the estimated exposure is 100 times lower than the NOAEL. The threshold for regulatory consideration agreed among Member States is an MOET of 100 at the 99.9th percentile of the exposure distribution.[1] MOETs below this threshold may therefore trigger risk management decisions by the European Commission and Member States.

It should be emphasised that results presented are exposure estimates based on the methods and assumptions listed in Section 2. These results do not estimate the actual risk of European consumers and do not account for all possible uncertainties. Although some uncertainties affecting the exposure estimates may already be highlighted in this report, the overall risk characterisation is addressed in a separate report (EFSA, 2019b). That report combines the assessment of all uncertainties related to both hazard assessment and exposure assessment into a consolidated risk characterisation.

3.1. Pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition (CAG-NAN)

3.1.1. Tier I

The results in Table 13 were obtained using the very conservative Tier I calculations. The largest margins of exposure were observed for adults, where MOET estimates at the 99.9th percentile ranged from 17.7 (Italian adults) to 31.2 (Czech adults). The margins of exposure for toddlers and other children were smaller. MOET estimates for these age classes ranged from 5.51 (Bulgarian children) to 15.3 (Dutch children).

The main drivers of exposure were identified for the upper percentile of the exposure distribution (see Annex B.1, Figure B.1.03 and Table B.1.02). Carbofuran made the greatest contribution to the upper-tail of the distribution (52–73%); most of this contribution came from tomatoes (up to 30%), cucumbers (up to 28%), mandarins (up to 22%), green beans with pods (up to 14%), peaches (up to 14%) and lettuces (up to 13%). Other important substances were triazophos (9–21%, green beans up to 17%), chlorpyrifos (2–8%, apples up to 6%), omethoate (2–7%, wine grapes up to 3%) and azinphos-ethyl (1–8%, peppers up to 7%). Remaining substances of the CAG only played a minor role in the upper-tail exposures (not more than 5%).

Although MOET estimates at the 99.9th percentile of the exposure distribution were all below 100, the Tier I calculations are by nature very conservative. This is clearly evidenced for CAG-NAN, where carbofuran was identified as the main driver of exposure. Carbofuran is part of an unspecific residue definition together with carbosulfan and benfuracarb (see Annex A.1, Table A.1.03). According to the Tier I assumptions, all measurements for unspecific residues are assigned to the most potent substance, i.e. carbofuran in this case. Assignment of all carbamate measurements to carbofuran therefore overestimates the exposure (i.e. underestimate MOET). Furthermore, most of the occurrence data for carbofuran are below the LOQ, with LOQ values going up to 0.5 mg/kg (see Annex A.1, Table A.1.10). As carbofuran is not authorised for use on any of the commodities under assessment (see Annex A.1, Table A.1.06), it is very unlikely that residues of carbofuran will be present in any of these samples. Yet, according to the Tier I assumptions, carbofuran was assumed to be present in all these samples at 1/2 LOQ.

These inaccuracies are accounted for under the Tier II assumptions.

Table 13: Estimates of the total margin of exposure (MOET) and their corresponding 95% confidence intervals at the 50th, 95th, 99th and 99.9th percentiles of the exposure distribution for the Tier I scenario of CAG-NAN

Country	Population class	50th Percentile	95th Percentile	99th Percentile	99.9th Percentile
Belgium	Adults	437 [426–447]	149 [144–153]	85.9 [80.8–91.6]	27.5 [22.5–32.8]
Czech Republic	Adults	479 [470–488]	149 [145–153]	83 [79.6–87.7]	31.2 [26.4–36.6]
Germany	Adults	393 [388–397]	130 [127–132]	73.2 [70.1–75.8]	22 [17.3–27.8]
Italy	Adults	351 [343–359]	120 [116–126]	64.3 [59.6–68.2]	17.7 [15.2–21.2]
Bulgaria	Other children	130 [125–135]	40.9 [39.1–42.7]	22 [20.8–23.8]	5.51 [4.47–7.07]
France	Other children	235 [229–240]	72.8 [70.5–74.7]	41 [38.7–43.3]	11.4 [9.36–13.5]
Netherlands	Other children	228 [222–234]	80.3 [77–84]	46.3 [43.8–49.3]	15.3 [12.7–17.9]
Denmark	Toddlers	161 [158–163]	55.3 [54–56.7]	31.7 [30.5–32.9]	9.84 [7.84–12.4]
Netherlands	Toddlers	176 [168–183]	59.4 [54.8–63.4]	34.1 [31–38.2]	12 [10.4–13.9]
United Kingdom	Toddlers	184 [181–187]	64.9 [63.1–66.5]	37.7 [35.8–39.6]	12 [10–15.3]

CAG-NAN: cumulative assessment group of pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition.

3.1.2. Tier II

The results from the Tier II calculations are displayed in Table 14. As with the Tier I results, the largest margins of exposure at the 99.9th percentile were observed for adults; adult MOETs ranged from 92.4 (Germany) to 121 (the Czech Republic). The MOETs for toddlers and other children ranged from 40.2 (Dutch toddlers) to 61.7 (UK toddlers).

The main drivers of exposures exceeding the 99th percentile were different from the Tier I calculations (see Annex C.1, Figure C.1.03 and Table C.1.02). Chlorpyrifos, triazophos and omethoate remained very important contributors, each accounting for, respectively, 17–48%, 11–53% and 7–26% of the upper-tail exposures. The commodities driving the contribution of chlorpyrifos were apples (up to 37%), wine grapes (up to 20%), peaches (up to 8%) and pears (up to 5%). The contribution of triazophos was mainly driven by green beans with pods (up to 50%) but peppers and rice also contributed substantially (up to 9%). For omethoate, the driving commodities were olives for oil production (14% for Italian adults) and tomatoes (up to 6%). While the contribution of carbofuran showed an important decline compared to Tier I (i.e. from 52–73% to 3–9%), new contributors under the Tier II assumptions are dichlorvos (2–8%) and formetanate (0.5–8%). The contribution of dichlorvos is almost exclusively attributed to its occurrence in wheat. Likewise, the contribution of formetante is mainly attributed to table grapes and strawberries.

The change in the main drivers of exposure was likely due to the different assumptions made in the Tier II calculations. Under this scenario, assignment of the active substances to unspecific residue definitions is no longer restricted to the most potent active substance. Furthermore, left-censored data are imputed with 1/2 LOQ based on authorised uses and estimated use frequencies. These two assumptions had a large impact on the exposure to carbofuran, which was the main driver of the exposure under Tier I. Overall, considering all changes in the assumptions, a 3- to 9-fold increase of the MOETs from Tier I to Tier II was observed.

Table 14: Estimates of the total margin of exposure (MOET) and their corresponding 95% confidence intervals at the 50th, 95th, 99th and 99.9th percentiles of the exposure distribution for the Tier II scenario of CAG-NAN

Country	Population class	50th Percentile	95th Percentile	99th Percentile	99.9th Percentile
Belgium	Adults	2,960 [2,880–3,030]	1,230 [1,160–1,270]	552 [432–618]	106 [72.1–153]
Czech Republic	Adults	2,870 [2,790–2,930]	1,210 [1,160–1,280]	584 [509–655]	121 [85.7–166]
Germany	Adults	2,330 [2,300–2,350]	987 [947–1,020]	414 [361–452]	92.4 [72.9–116]
Italy	Adults	3,820 [3,540–4,030]	1,090 [808–1,320]	390 [266–545]	96.5 [70.9–131]
Bulgaria	Other children	1,970 [1,890–2,070]	606 [580–639]	236 [203–269]	48.6 [36.5–63.4]
France	Other children	2,570 [2,490–2,660]	782 [733–826]	287 [253–321]	59.1 [43–72.7]
Netherlands	Other children	2,170 [2,110–2,220]	617 [588–660]	229 [200–254]	52 [41.9–62]
Denmark	Toddlers	1,630 [1,600–1,660]	499 [475–522]	211 [193–229]	58.9 [45.3–72.8]
Netherlands	Toddlers	1,670 [1,580–1,760]	471 [438–501]	173 [154–187]	40.2 [32.7–47]
United Kingdom	Toddlers	1,660 [1,610–1,700]	592 [573–617]	261 [231–285]	61.7 [44.7–74]

CAG-NAN: cumulative assessment group of pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition.

Although Tier II calculations are expected to reflect a more refined scenario, this scenario is still subject to uncertainties. Some of these uncertainties were addressed through sensitivity analyses. A comparison between the MOETs obtained at the 99.9th percentile from the Tier II calculations and their corresponding sensitivity analyses is made in Table 15.

Sensitivity analyses A and B tested the uncertainty of imputing left-censored data with 1/2 LOQ based on an estimation of use frequencies. Sensitivity analysis A imputes all left-censored data with 1/2 LOQ based on authorisation rather than use frequency. This is over-conservative, as the commodities are not expected to be treated with all authorised substances at the same time. When applying these assumptions, the MOETs dropped by 1.2–1.4 times. On the other hand, sensitivity analysis B imputes all left-censored data with zero. This may not be sufficiently conservative, as not all left-censored data would be true zeros. Nevertheless, the MOETs remained steady in this case and minor variations were mainly attributed to the random effect of probabilistic modelling. This outcome may seem unexpected because, on a theoretical basis, the imputation based on use frequency would be more conservative than sensitivity analysis B. In the case of chlorpyrifos for example (i.e. the main drivers for the acute exposure), it was estimated under Tier II that over 80% of the wine grapes, oranges and carrots were treated with chlorpyrifos (see Annex C.1, Table C.1.05). Regardless of these conservative assumptions, however, the imputation of left-censored data has a limited impact on the outcome of the assessment. This finding is consistent with the expectation that MOETs at the 99.9th percentile of the acute exposure distribution are primarily driven by samples with quantifiable findings.

Sensitivity analysis C investigated the effect that missing processing factors might have on the margins of exposure. When no residues were assumed to be present in processed foods, the MOETs rose by a factor of 1.6–2.5. This change indicates that processing factors were not available for important risk drivers, such as chlorpyrifos in apple juice and wine, and omethoate in olive oil. Although including these processing factors might not increase the margins of exposure to the extent suggested in the sensitivity analysis, they would contribute to a substantial reduction of the uncertainty. In addition, it is noted that most consumption data for green beans were reported without specifying any preparation method. It was therefore not possible, within this sensitivity test, to account for the potential loss of triazophos during cooking of green beans. Based on the high contribution of triazophos in green beans to the exposure, this is also expected to overestimate the exposure.

Sensitivity analysis D investigated the effect of excluding foods for infants and young children in CAG-NAN. There were no substantial changes in the margins of exposure when this assumption was made. This confirms previous findings of EFSA that exposure of toddlers to pesticide residues mainly comes from conventional foods (EFSA PPR Panel, 2018). This is due to the low levels of pesticide residues found in foods for infants and young children.

Table 15: Estimates of the total margin of exposure (MOET) and their corresponding 95% confidence intervals at the 99.9th percentiles of the exposure distribution for the Tier II scenario of CAG-NAN and several sensitivity analyses

Country	Population class	Tier II	Sensitivity analysis A[(a)]	Sensitivity analysis B[(b)]	Sensitivity analysis C[(c)]	Sensitivity analysis D[(d)]
Belgium	Adults	106 [72.1–153]	87.9 [64.2–108]	103 [65.5–154]	261 [212–308]	102 [76.3–149]
Czech Republic	Adults	121 [85.7–166]	97.2 [74.8–118]	119 [84.1–172]	296 [259–343]	122 [92.3–167]
Germany	Adults	92.4 [72.9–116]	74.1 [60.3–84.6]	93 [71.6–120]	222 [192–260]	93.4 [75.6–124]
Italy	Adults	96.5 [70.9–131]	81.7 [62.8–100]	96.3 [71.2–123]	184 [151–230]	97.5 [75.6–128]
Bulgaria	Other children	48.6 [36.5–63.4]	37.9 [31.5–44.7]	49.2 [34.7–61.9]	92.1 [71.6–111]	48.1 [36–60.4]
France	Other children	59.1 [43–72.7]	47.6 [39.9–57.8]	59.9 [45.6–80]	94 [66–126]	58.9 [44.5–74.1]
Netherlands	Other children	52 [41.9–62]	42.7 [36.3–50.1]	52.3 [42.4–62.1]	112 [91.2–139]	52.9 [41.4–65]
Denmark	Toddlers	58.9 [45.3–72.8]	41.8 [34.7–48]	60.6 [41.9–76.4]	99.6 [79–115]	58.6 [42.7–73.8]
Netherlands	Toddlers	40.2 [32.7–47]	30.9 [26.1–36.9]	40.2 [32.1–50.4]	91.1 [73.2–111]	38.8 [30.6–48.2]
United Kingdom	Toddlers	61.7 [44.7–74]	46.3 [37.5–51.8]	60 [45.9–73.9]	138 [111–168]	59.6 [45.1–76.9]

CAG-NAN: cumulative assessment group of pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition.
(a): Sensitivity analysis assuming that left-censored data are at 1/2 LOQ when the use of the active substance is authorised.
(b): Sensitivity analysis assuming that all left-censored data are at zero.
(c): Sensitivity analysis assuming that residues will not be present in any processed food.
(d): Sensitivity analysis excluding foods for infants and young children.

An additional uncertainty arises from the occurrence data that are missing for certain combinations of active substances and commodities. This uncertainty, however, cannot be quantified by means of a sensitivity analysis. For CAG-NAN, the missing occurrence data mainly refer to combinations where the use of the active substance is not authorised, except for the use of ethephon in olives for oil production (see Annex C.1, Table C.1.04). Considering that ethephon has a relatively high NOAEL within the CAG (see Annex A.1, Table A.1.01), this uncertainty is not expected to have a substantial impact on the outcome of the exposure assessment.

Apart from the uncertainties, it is also important to understand the factors that influence the lowest MOETs. Individual days with a MOET lower than the MOET calculated at the 99th percentile of the exposure distribution were therefore extracted for the nominal run of each dietary survey (see Annex C.1, Table C.1.03). Analysis of these individual records revealed that, for toddlers and children, 75% of the extracted individual days was driven by single substance-commodity combinations, each contributing for more than 80% of the total exposure within that day. Around 95% of these substance-commodity combinations were found to exceed the MRL that was in place at the end of the reference period (i.e. end of 2016). For some of these substance-commodity combinations, MRLs were lowered shortly before the end of the reference period (e.g. MRLs for chlorpyrifos in apples and pears were lowered in August 2016). This means that some samples, despite being compliant with the MRL at the beginning of the reference period, would no longer be placed on the market when the latest legal standards were to be applied. It is therefore expected that exposure from these combinations will decrease after the reference period.

Although a more detailed analysis would be recommended for the future, the above findings indicate that the highest acute exposures do not result from cumulative exposure to multiple

compounds. Furthermore, acute exposures are strongly influenced by the presence of products on the market that are not compliant with the latest legal standards.

3.2. Pesticides associated with functional alterations of the motor division (CAG-NAM)

3.2.1. Tier I

Table 16 displays the results from the Tier I calculations. Similar to the CAG-NAN, the largest MOETs were observed in adults, where MOET estimates at the 99.9th percentile were between 48.3 (Italy) to 61.3 (Belgium). Likewise, the lowest MOETs were in toddlers and other children. MOET estimates in these population classes ranged from 16.9 (Bulgarian children) to 29.1 (Dutch children).

The main contributors were identified for consumers with an exposure exceeding the 99th percentile of the distribution (see Annex B.2, Figure B.2.03 and Table B.2.02). Triazophos was the predominant contributor to the exposure (15–35%), most of this coming from cucumbers (up to 20%), beans with pods (up to 19%), aubergines (up to 11%) and sweet peppers (up to 10%). In addition, thiram (13–22%) and azinphos-ethyl (3–18%) contributed substantially to the exposure. Most thiram came from wheat (up to 9%), apples (up to 7%) and wine grapes (up to 5%) whilst most azinphos-ethyl came from sweet peppers (up to 16%) and cauliflower (7%). Any other active substance contributed less than 10% and remaining substance-commodity combinations never exceeded 4%.

While all the MOET estimates at the 99.9th percentile were below 100, several very conservative assumptions were made to generate these results. As with carbofuran in the CAG-NAN, one of the main contributors, thiram, is the most potent substance in the dithiocarbamates group (see Annex A.2, Table A.2.03). According to Tier I assumptions, all measurements for dithiocarbamates are assigned to thiram, although less potent dithiocarbamates might be present in practice. Thiram is also not authorised for use in wheat. Assigning thiram to this commodity is therefore even more conservative. Furthermore, the contribution of azinphos-ethyl is driven by the high number of left-censored data (99.8–100.0%). Under Tier I assumptions all these results were assumed to be 1/2 LOQ. This is most likely conservative considering that azinphos-ethyl is not authorised for use on any crop (see Annex A.2, Table A.2.06). These overestimations were accounted for in the Tier II calculations.

Table 16: Estimates of the total margin of exposure (MOET) and their corresponding 95% confidence intervals at the 50th, 95th, 99th and 99.9th percentiles of the exposure distribution for the Tier I scenario of CAG-NAM

Country	Population class	50th Percentile	95th Percentile	99th Percentile	99.9th Percentile
Belgium	Adults	560 [550–571]	213 [207–221]	130 [121–137]	61.3 [52.4–70.3]
Czech Republic	Adults	584 [575–595]	201 [195–208]	113 [107–121]	54.2 [48.2–62.3]
Germany	Adults	493 [489–499]	181 [177–184]	105 [101–109]	48.7 [44.5–53.8]
Italy	Adults	505 [492–519]	187 [174–194]	104 [95.1–112]	48.3 [42.9–55.8]
Bulgaria	Other children	184 [178–190]	61.2 [58–64.5]	34.9 [31.5–37.9]	16.9 [15–18.7]
France	Other children	281 [272–287]	98 [94.7–100]	60.1 [56.6–63.3]	28.6 [24.3–31.9]
Netherlands	Other children	258 [253–265]	97.1 [93.6–101]	59.4 [56.3–63.2]	29.1 [26–32.8]
Denmark	Toddlers	187 [184–190]	72.2 [71–74]	44.9 [42.1–47.5]	22.9 [20.7–24.7]
Netherlands	Toddlers	198 [190–209]	71.5 [67.1–76.2]	43.8 [39.3–48.4]	20.9 [17.8–24.7]
United Kingdom	Toddlers	210 [207–214]	86 [83.7–87.8]	53.9 [51.3–56.4]	26.3 [23.6–29.2]

CAG-NAM: cumulative assessment group of pesticides associated with functional alterations of the motor division.

3.2.2. Tier II

Table 17 lists the results from the Tier II calculations. The largest margins of exposure at the 99.9th percentile were still observed in adults. The MOET estimates ranged from 145 (Italy) to 181 (Belgium). The lowest margins of exposure were seen in toddlers and other children. These ranged from 66.3 (Bulgarian children) to 89.5 (Dutch children).

The main drivers of the exposure for the highest percentile of the distribution were different in Tier II compared to Tier I (see Annex C.2, Figure C.2.03 and Table C.2.02). Although triazophos and thiram remained predominant contributors to the exposure in Tier II (13–57% and 2–17%, respectively), the food commodities affecting these contributions were different. For triazophos, the contribution was no longer driven by its occurrence in beans with pods and sweet peppers only (up to 50% and 5%). Rice was now also identified as an important contributor (up to 14%). The contribution of thiram on the other hand came from wine grapes (up to 6%) and apples (up to 6%). Furthermore, while the contribution of azinphos-ethyl declined to less than 4% in all surveys, the contribution of deltamethrin increased to 9–21%. This contribution was mainly attributed to the presence of deltamethrin in wheat and oats (up to 19% and 9%, respectively). Other important contributions came from chlormequat in oats (up to 17%), beta-cypermethrin in wheat (up to 12%), omethoate in olives for oil productions (up to 11%) and acrinathrin in bananas (up to 6%). The remaining substance-commodity combinations did not contribute to more than 5% of the exposure in any dietary survey.

As with the CAG-NAN, the change in the assumptions made in the Tier II calculations resulted in a substantial increase of the MOETs. In this case, however, a smaller increase of the MOETs from Tier I to Tier II was observed, i.e. a three- to fourfold increase for CAG-NAM compared to a three- to ninefold increase for CAG-NAN.

Table 17: Estimates of the total margin of exposure (MOET) and their corresponding 95% confidence intervals at the 50th, 95th, 99th and 99.9th percentiles of the exposure distribution for the Tier II scenario of CAG-NAM

Country	Population class	50th Percentile	95th Percentile	99th Percentile	99.9th Percentile
Belgium	Adults	3,380 [3,290–3,460]	1,350 [1,260–1,420]	651 [523–744]	180 [122–259]
Czech Republic	Adults	3,180 [3,100–3,240]	1,330 [1,260–1,400]	635 [551–733]	181 [133–243]
Germany	Adults	2,630 [2,600–2,670]	1,130 [1,080–1,170]	574 [502–643]	170 [128–216]
Italy	Adults	4,330 [4,060–4,610]	1,210 [922–1480]	475 [379–638]	145 [102–190]
Bulgaria	Other children	2,120 [2,030–2,240]	645 [607–678]	249 [196–285]	66.3 [51.4–93.5]
France	Other children	2,750 [2,680–2,840]	792 [735–850]	318 [271–362]	84.3 [65.6–111]
Netherlands	Other children	2,340 [2,280–2,410]	737 [690–775]	305 [240–349]	89.5 [71.7–113]
Denmark	Toddlers	1,640 [1,570–1,710]	450 [388–528]	212 [179–257]	80.7 [64.1–101]
Netherlands	Toddlers	1,850 [1,770–1,940]	588 [558–638]	243 [206–274]	69.5 [56–88.4]
United Kingdom	Toddlers	1,790 [1,740–1,830]	574 [540–621]	244 [208–277]	74.4 [59.4–90.9]

CAG-NAM: cumulative assessment group of pesticides associated with functional alterations of the motor division.

For CAG-NAM, uncertainties were addressed by means of the same sensitivity analyses as CAG-NAN. The comparison between the MOETs obtained at the 99.9th percentile of the Tier II exposure distribution and the sensitivity analyses is reported in Table 18.

When left-censored data are imputed with 1/2 LOQ based on authorisation rather than use frequency (sensitivity analysis A), the MOETs dropped by 1.2–1.7 times. When all left-censored data are imputed with zero (sensitivity analysis B), only minor variations are observed compared to Tier II.

As for CAG-NAN, these findings confirm that MOETs at the 99.9th percentile of the acute exposure distribution are primarily driven by samples with quantifiable findings.

When no residues were assumed to be present in processed foods (sensitivity analysis C), the MOETs rose by a factor of 1.9–3.5. This increase of MOETs is slightly larger than the increase calculated for CAG-NAN. Most likely this is due to the important contribution of cereals which are consumed processed and for which limited processing factors are available. Furthermore, processing factors were not available for other important risk drivers, such as thiram in wine, omethoate in olive oil and triazophos in green beans. Processing factors for these commodities are therefore expected to substantially reduce the uncertainty.

Sensitivity analysis D confirmed previous findings that exposure of toddlers to pesticide residues mainly comes from conventional foods.

Table 18: Estimates of the total margin of exposure (MOET) and their corresponding 95% confidence intervals at the 99.9th percentiles of the exposure distribution for the Tier II scenario of CAG-NAM and several sensitivity analyses

Country	Population class	Tier II	Sensitivity analysis A[a]	Sensitivity analysis B[b]	Sensitivity analysis C[c]	Sensitivity analysis D[d]
Belgium	Adults	180 [122–259]	128 [104–152]	181 [133–266]	441 [345–536]	192 [116–251]
Czech Republic	Adults	181 [133–243]	126 [107–146]	179 [129–253]	520 [420–620]	186 [134–250]
Germany	Adults	170 [128–216]	110 [93.7–124]	174 [127–230]	319 [239–420]	178 [122–213]
Italy	Adults	145 [102–190]	119 [94.7–140]	147 [107–205]	297 [225–357]	149 [112–208]
Bulgaria	Other children	66.3 [51.4–93.5]	47.2 [40.6–53.5]	67.5 [52.4–90.6]	155 [115–189]	69.5 [51.5–82.8]
France	Other children	84.3 [65.6–111]	65.2 [54.1–74.5]	86.5 [68.3–121]	170 [102–239]	88.2 [68.4–108]
Netherlands	Other children	89.5 [71.7–113]	64.3 [54.7–71]	89.9 [72.1–120]	222 [155–289]	89.7 [75.3–120]
Denmark	Toddlers	80.7 [64.1–101]	50 [43.4–55.2]	79.2 [64.3–101]	181 [144–221]	80.7 [67.3–100]
Netherlands	Toddlers	69.5 [56–88.4]	41.8 [36.3–51.7]	70.8 [53.1–88.8]	179 [147–223]	69.7 [56.2–86.2]
United Kingdom	Toddlers	74.4 [59.4–90.9]	54.3 [46.8–59.5]	74.1 [58–94]	263 [209–303]	74.4 [63.2–89.7]

CAG-NAM: cumulative assessment group of pesticides associated with functional alterations of the motor division.
(a): Sensitivity analysis assuming that left-censored data are at 1/2 LOQ when the use of the active substance is authorised.
(b): Sensitivity analysis assuming that all left-censored data are at zero.
(c): Sensitivity analysis assuming that residues will not be present in any processed food.
(d): Sensitivity analysis excluding foods for infants and young children.

As for CAG-NAN, additional uncertainty arises from the occurrence data that are missing for certain combinations of active substances and commodities. In this case, several combinations referred to authorised uses of penflufen, glufosinate and ethephon (see Annex C.2, Table C.2.04). Considering, however, that all three substances have a relatively high NOAEL within the CAG (see Annex A.2, Table A.2.01), this uncertainty is not expected to have a substantial impact on the outcome of the exposure assessment.

Apart from the uncertainties, factors influencing the lowest MOETs were also investigated. In this case, the analysis revealed that, for toddlers and children, 78% of the extracted individual days was driven by single substance-commodity combinations, each contributing for more than 80% of the total exposure within that day. Around 40% of these substance-commodity combinations were found to exceed the MRL that was in place at the end of the reference period (i.e. end of 2016). These findings confirm the previous finding that the highest acute exposures do not result from cumulative exposure to multiple compounds. These exposures are also in this case influenced by the presence of samples that are not compliant with the latest legal standards but to a smaller extend compared to CAG-NAN.

3.3. Comparison with Monte Carlo Risk Assessment (MCRA) software

Results of the calculations obtained by EFSA were compared with results obtained by the RIVM using the MCRA software version 8.3 (van Klaveren et al., 2019). Results obtained with both software are presented in Table 19 and Table 20 for the Tier II scenarios of CAG-NAN and CAG-NAM, respectively. This comparison is presented for the higher percentiles of the exposure distribution because these percentiles are considered most relevant for risk management purposes.

Table 19: Estimates of the total margin of exposure (MOET) and their corresponding 95% confidence intervals obtained with two different software (SAS® and MCRA) at the 99th and 99.9th percentiles of the exposure distribution for the Tier II scenario of CAG-NAN

Country	Population class	SAS®		MCRA	
		99th Percentile	99.9th Percentile	99th Percentile	99.9th Percentile
Belgium	Adults	552 [432–618]	106 [72.1–153]	514 [420–607]	102 [72–162]
Czech Republic	Adults	584 [509–655]	121 [85.7–166]	542 [462–648]	120 [87–176]
Germany	Adults	414 [361–452]	92.4 [72.9–116]	413 [362–463]	95 [73–120]
Italy	Adults	390 [266–545]	96.5 [70.9–131]	331 [263–537]	96 [75–149]
Bulgaria	Other children	236 [203–269]	48.6 [36.5–63.4]	241 [210–264]	49 [36–63]
France	Other children	287 [253–321]	59.1 [43–72.7]	271 [221–320]	59 [46–74]
Netherlands	Other children	229 [200–254]	52 [41.9–62]	227 [202–249]	52 [45–62]
Denmark	Toddlers	211 [193–229]	58.9 [45.3–72.8]	213 [191–227]	60 [50–69]
Netherlands	Toddlers	173 [154–187]	40.2 [32.7–47]	169 [152–188]	40 [33–50]
United Kingdom	Toddlers	261 [231–285]	61.7 [44.7–74]	258 [227–287]	61 [47–76]

CAG-NAN: cumulative assessment group of pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition.

Table 20: Estimates of the total margin of exposure (MOET) and their corresponding 95% confidence intervals obtained with two different software (SAS® and MCRA) at the 99th and 99.9th percentiles of the exposure distribution for the Tier II scenario of CAG-NAM

Country	Population class	SAS®		MCRA	
		99th Percentile	99.9th Percentile	99th Percentile	99.9th Percentile
Belgium	Adults	651 [523–744]	180 [122–259]	629 [535–717]	176 [115–228]
Czech Republic	Adults	635 [551–733]	181 [133–243]	614 [544–700]	172 [131–236]
Germany	Adults	574 [502–643]	170 [128–216]	578 [517–640]	171 [127–211]
Italy	Adults	475 [379–638]	145 [102–190]	461 [379–588]	141 [109–185]
Bulgaria	Other children	249 [196–285]	66.3 [51.4–93.5]	238 [195–274]	63 [53–81]
France	Other children	318 [271–362]	84.3 [65.6–111]	303 [270–353]	84 [65–102]

Country	Population class	SAS®		MCRA	
		99th Percentile	99.9th Percentile	99th Percentile	99.9th Percentile
Netherlands	Other children	305 [240–349]	89.5 [71.7–113]	303 [255–351]	89 [75–111]
Denmark	Toddlers	212 [179–257]	80.7 [64.1–101]	217 [189–248]	80 [63–100]
Netherlands	Toddlers	243 [206–274]	69.5 [56–88.4]	236 [206–266]	68 [56–85]
United Kingdom	Toddlers	244 [208–277]	74.4 [59.4–90.9]	245 [218–272]	73 [61–89]

CAG-NAM: cumulative assessment group of pesticides associated with functional alterations of the motor division.

The results obtained with SAS® and MCRA were found to be very similar. For CAG-NAM, most estimates calculated by EFSA showed less than 5% deviation compared to the estimates obtained by RIVM. For CAG-NAN, the observed deviations generally remained below 8%. One dietary survey (Italian adults) showed a more substantial deviation of 15% on the estimate but the confidence intervals calculated for this dietary survey were nearly identical. These minor divergencies are therefore attributed to random effects which are inherent to the probabilistic methodologies applied in the exposure assessment.

Despite obtaining very similar results at the 99th and 99.9th percentiles of the exposure distributions, any potential cause for variances was further investigated by EFSA and RIVM. As a result, minor methodological differences were identified and summarised as follows:

- Within the SAS® program the nominal run (i.e. the run based on the original data) was considered equivalent to any of the bootstrap runs, and uncertainty percentiles were based on all runs. In MCRA, uncertainty percentiles are based on the bootstrap runs only. Theoretically, this may cause a difference because two randomly selected bootstrapped data sets will on average differ more from each other than each of them will differ from the original data set. However, when the calculations are performed with one nominal run and 99 bootstrap runs (SAS) or with 100 bootstrap runs (MCRA), the impact of this difference is anticipated to be very small in practice.
- To reduce computational time, calculations are handled through parallel computation. In MCRA, parallel computation is used within each bootstrap iteration, and bootstraps are handled sequentially. In the SAS® program, however, it was decided to handle the bootstraps in parallel. This approach caused a bias in the random assignment of active substances to unspecific residue definitions for the bootstrapped data (not for the nominal run), which could not be resolved by EFSA. Hence, the confidence interval calculated with SAS is biased (i.e. too wide) when the exposure is driven by substances coming from an unspecific residue definition.

Based on the above considerations it is concluded that MCRA and the SAS® program elaborated by EFSA produce nearly identical results at the percentiles of interest. The differences observed between both software are mainly attributed to the random effect of probabilistic modelling. Yet, these differences do not impact on the outcome of the assessment.

4. Conclusions

EFSA performed a retrospective assessment of cumulative exposure to pesticides affecting the nervous system for the reference period 2014–2016. Calculations included 10 dietary surveys that cover different age classes and geographical areas. In addition, EFSA only considered the food consumption data for water, foods for infants and young children and 30 RPCs of plant origin that are widely consumed within Europe. All calculations were successfully executed using SAS® software.

As agreed by risk managers in the SC PAFF, this methodology is applied in a tiered approach. While the Tier I scenario uses very conservative assumptions for an efficient screening of the exposure with low risk for underestimation, the Tier II scenario includes assumptions that are more refined but still intended to be conservative. For each scenario, exposure estimates were obtained for different percentiles of the exposure distribution and the total margin of exposure (MOET, i.e. the ratio of the toxicological reference dose to the estimated exposure) was calculated at each percentile. In

accordance with the threshold agreed at the SC PAFF, further regulatory consideration would be required when the MOET calculated at the 99.9th percentile of the exposure distribution is below 100.

The lowest MOET estimates were obtained for pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition. According to the Tier II scenario, MOET estimates at the 50th, 95th and 99th percentile of the exposure distribution were all well above 100. At the 99.9th percentile, estimates were mostly below 100, ranging from 40.2 to 61.7 in toddlers and other children. For adults, the MOETs were ranging from 92.4 to 121. The exposure to this group of pesticides was predominantly driven by the occurrence of chlorpyrifos, triazophos and omethoate. Other important contributors included dichlorvos, formetanate and carbofuran.

For pesticides associated with functional alterations of the motor division, MOETs calculated at the 99.9th percentile of the exposure distribution were higher, ranging from 66.3 to 89.5 in toddlers and other children and from 145 to 181 in adults. In this case, the main drivers for the exposure were identified as triazophos and deltamethrin. Other important contributions came from beta-cypermethrin, omethoate, thiram, chlormequat and acrinathrin.

To ensure a rigorous validation of the methodology, exposure estimates obtained by EFSA were validated against those obtained by the RIVM using the MCRA software, version 8.3. Comparison of the results revealed that both tools produced nearly identical results and any observed differences are mainly attributed to the random effect of probabilistic modelling. These minor differences, however, do not impact on the outcome of the exposure assessment.

It is important to note that all calculations were conducted with conservative assumptions likely to overestimate the exposure, even in the more refined Tier II scenario. The most important overestimation originates from the limited availability of processing factor data. When such data are missing, it is assumed that all pesticides in the RPC will reach the end consumer without any loss of residues. Sensitivity tests have demonstrated that, for pesticides associated with functional alterations of the motor division, further data on the effect of processing might result in a threefold increase of MOET estimates in toddlers. Information on the effect of processing prior to consumption is therefore expected to substantially reduce the uncertainty, particularly in the case of risk drivers.

Another impactful assumption is the random assignment of active substances to unspecific residue definitions (i.e. a measurement that may comprise multiple active substances). All substances were assumed to have an equal probability of being applied while this is not necessarily reflecting the actual use of the pesticides. Comparison with the Tier I scenario has indicated that this assumption is very impactful. If data on the use frequency of pesticides would be available at EU level, this assumption could be further refined.

Uncertainties caused by the imputation of left-censored data (i.e. measurements below the limit of quantification), were found to have a very limited impact. This finding is consistent with the expectation that MOETs at the 99.9th percentile of the acute exposure distribution are primarily driven by samples with quantifiable findings.

Analysis of the factors influencing the upper part of the exposure distribution also revealed that the highest acute exposures do not result from cumulative exposure to multiple compounds. For toddlers and children, around 75% of the high exposure estimates were driven by a single substance in a specific commodity. For pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition, around 95% of these combinations were found to exceed the MRL that was in place at the end of the reference period. It should be noted, however, that MRLs for some of these combinations were lowered shortly before the end of the reference period (e.g. MRLs for chlorpyrifos in apples and pears were lowered in August 2016). It is therefore expected that, in the meantime, exposure from these combinations will have decreased already.

Uncertainties considered in this assessment, however, only refer to the exposure calculations and should still be considered in conjunction with other uncertainties that may apply to the hazard characterisation. Hence, together with the results obtained by RIVM, the exposure estimates presented in this report are used for the final scientific report on the cumulative dietary risk characterisation of pesticides that have acute effects on the nervous system. The latter combines the hazard assessment and exposure assessment into a consolidated risk characterisation including all related uncertainties.

5.   Recommendation

In view of refining the exposure estimates, the following recommendations were derived by EFSA.

- The assignment of active substances to unspecific residue definitions relies on the assumption that active substances are used with equal probability. Considering that this assumption has a

strong impact on the exposure estimates, it is recommended that EFSA, the European Commission and Member States collect data on the actual usage of pesticides.

- Due to the limited availability of processing factors for pesticide residues in food, the effect of processing prior to consumption is not adequately addressed in the current calculations. It is therefore recommended that EFSA, the European Commission and Member States further elaborate the European database on processing factors.

References

Efron B and Tibshirani R, 1993. An Introduction to the Bootstrap. Chapman & Hall.

EFSA (European Food Safety Authority), 2010. Guidance of EFSA on the standard sample description in food and feed. EFSA Journal 2010;8(1):1457, 54 pp. https://doi.org/10.2903/j.efsa.2010.1457

EFSA (European Food Safety Authority), 2011. Guidance of EFSA on the use of the EFSA Comprehensive European Food Consumption Database in Intakes Assessment. EFSA Journal 2011;9(3):2097, 34 pp. https://doi.org/10.2903/j.efsa.2011.2097

EFSA (European Food Safety Authority), 2013. Guidance of EFSA on the standard sample description ver. 2.0. EFSA Journal 2013;11(10):3424, 114 pp. https://doi.org/10.2903/j.efsa.2013.3424

EFSA (European Food Safety Authority), 2015a. Scientific report on the pesticide monitoring program: design assessment. EFSA Journal 2015;13(2):4005, 52 pp. https://doi.org/10.2903/j.efsa.2015.4005

EFSA (European Food Safety Authority), 2015b. The food classification and description system FoodEx2 (revision2). EFSA Supporting Publication 2015:EN-804, 90 pp. https://doi.org/10.2903/sp.efsa.2015.en-804

EFSA (European Food Safety Authority), 2016. The 2014 European Union report on pesticide residues in food. EFSA Journal 2016;14(10):4611, 139 pp. https://doi.org/10.2903/j.efsa.2016.4611

EFSA (European Food Safety Authority), 2017. The 2015 European Union report on pesticide residues in food. EFSA Journal 2017;15(4):4791, 134 pp. https://doi.org/10.2903/j.efsa.2017.4791

EFSA (European Food Safety Authority), 2018a. The 2016 European Union report on pesticide residues in food. EFSA Journal 2018;16(7):5348, 139 pp. https://doi.org/10.2903/j.efsa.2018.5348

EFSA (European Food Safety Authority), 2018b. PRIMo rev 3 - Pesticide Residue Intake Model calculator (Version 3). Zenodo, https://doi.org/10.5281/zenodo.1137758

EFSA (European Food Safety Authority), 2019a. Scientific report on the establishment of cumulative assessment groups of pesticides for their effects on the nervous system. EFSA Journal 2019;17(9):5800, 125 pp. https://doi.org/10.2903/j.efsa.2019.5800

EFSA (European Food Safety Authority), 2019b. Scientific report on the cumulative dietary risk characterisation of pesticides that have acute effects on the nervous system (draft submitted for public consultation on 17 September 2019). Available online: https://www.efsa.europa.eu/en/consultations/call/public-consultation-scientific-report-cumulative

EFSA (European Food Safety Authority), 2019c. Harmonized terminology for scientific research [Data set]. Zenodo. https://doi.org/10.5281/zenodo.2554064

EFSA (European Food Safety Authority), 2019d. Technical report on the raw primary commodity (RPC) model: strengthening EFSA's capacity to assess dietary exposure at different levels of the food chain, from raw primary commodities to foods as consumed. EFSA Supporting Publication 2019:EN-1532, 30 pp. https://doi.org/10.2903/sp.efsa.2019.en-1532

EFSA PPR Panel (European Food Safety Authority Panel on Plant Protection Products and their Residues), 2008. Scientific Opinion on 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. EFSA Journal 2008;6(5):704, 85 pp. https://doi.org/10.2903/j.efsa.2008.704

EFSA PPR Panel (European Food Safety Authority Panel on Plant Protection Products and their Residues), 2009. 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. EFSA Journal 2009;7(9):1167, 104 pp. https://doi.org/10.2903/j.efsa.2009.1167

EFSA PPR Panel (European Food Safety Authority Panel on Plant Protection Products and their Residues), 2012. Guidance on the use of probabilistic methodology for modelling dietary exposure to pesticide residues. EFSA Journal 2012;10(10):2839, 95 pp. https://doi.org/10.2903/j.efsa.2012.2839

EFSA PPR Panel (European Food Safety Authority Panel on Plant Protection Products and their Residues), 2013a. Scientific Opinion on the identification of pesticides to be included in cumulative assessment groups on the basis of their toxicological profile. EFSA Journal 2013;11(7):3293, 131 pp. https://doi.org/10.2903/j.efsa.2013.3293

EFSA PPR Panel (European Food Safety Authority Panel on Plant Protection Products and their Residues), 2013b. Scientific Opinion on relevance of dissimilar mode of action and its appropriate application for cumulative risk assessment of pesticides residues in food. EFSA Journal 2013;11(12):3472, 40 pp. https://doi.org/10.2903/j.efsa.2013.3472

EFSA PPR Panel (European Food Safety Authority Panel on Plant Protection Products and their Residues), 2018. Scientific opinion on pesticides in foods for infants and young children. EFSA Journal 2018;16(6):5286, 75 pp. https://doi.org/10.2903/j.efsa.2016.5286

European Commission, 2017.Appendix D. Guidelines on comparability, extrapolation, group tolerances and data requirements for setting MRLs. 7525/VI/95-rev.10.3, June 2017.

FAO (Food and Agriculture Organization of the United Nations), 2003. Pesticide residues in food – 2005. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Expert Group on Pesticide Residues. FAO Plant Production and Protection Paper 176.

Huybrechts I, Sioen I, Boon PE, Ruprich J, Lafay L, Turrini A, Amiano P, Hirvonen T, De Neve M, Arcella D, Moschandreas J, Westerlund A, Ribas-Barba L, Hilbig A, Papoutsou S, Christensen T, Oltarzewski M, Virtanen S, Rehurkova I, Azpiri M, Sette S, Kersting M, Walkiewicz A, Serra-Majem L, Volatier J-L, Trolle E, Tornaritis M, Busk L, Kafatos A, Fabiansson S, De Henauw S and Van Klaveren JD, 2011. Dietary exposure assessments for children in Europe (the EXPOCHI project): rationale, methods and design. Archives of Public Health, 69, 4. https://doi.org/10.1186/0778-7367-69-4

van Klaveren JD, Kruisselbrink JW, de Boer WJ, van Donkersgoed G, te Biesebeek JD, Sam M and van der Voet H 2019. Cumulative dietary exposure assessment of pesticides that have acute effects on the nervous system using MCRA software. EFSA supporting publication 2019:EN-1708. 98 pp. https://doi.org/10.2903/sp.efsa.2019.en-1708

Kruisselbrink JW, van der Voet H, van Donkersgoed G and van Klaveren J, 2018. Proposal for a data model for probabilistic cumulative dietary exposure assessments of pesticides in line with the MCRA software. EFSA supporting publication 2018:EN-1375, 35 pp. https://doi.org/10.2903/sp.efsa.2018.en-1375

Merten C, Ferrari P, Bakker M, Boss A, Hearty A, Leclercq C, Lindtner O, Tlustos C, Verger P, Volatier JL and Arcella D, 2011. Methodological characteristics of the national dietary surveys carried out in the European Union as included in the European Food Safety Authority (EFSA) Comprehensive European Food Consumption Database. Food Additives and Contaminants: Part A, 28, 975–995. https://doi.org/10.1080/19440049.2011.576440.

Scholz R, van Donkersgoed G, Hermann M, Kittelmann A, von Schledorn M, Graven C, Mahieu K, van der Velde-Koerts T, Anagnostopoulos C, Bempelou E and Michalski B, 2018. Database of processing techniques and processing factors compatible with the EFSA food classification and description system FoodEx 2. Objective 3: European database of processing factors for pesticides in food. EFSA supporting publication 2018:EN-1510, 50 pp. https://doi.org/10.2903/sp.efsa.2018.en-1510

van der Voet H, de Boer W J, Kruisselbrink JW, van Donkersgoed G and van Klaveren J, 2016. MCRA made scalable for large cumulative assessment groups. EFSA supporting publication 2016:EN-910, 38 pp. https://doi.org/10.2903/sp.efsa.2016.en-910

Abbreviations

AS	active substance
AUP	agricultural use pattern, i.e. the combination of pesticide uses applied to a single commodity or crop
BW	body weight
CAG	cumulative assessment group
CAG-NAM	cumulative assessment group of pesticides associated with functional alterations of the motor division
CAG-NAN	cumulative assessment group of pesticides associated with brain and/or erythrocyte acetylcholinesterase inhibition
EUCP	EU-coordinated control programme
FP7	7th Framework Programme for Research
FoodEx2	multipurpose food classification and description system developed by EFSA, revision 2
LOQ	limit of quantification
MCRA	Monte Carlo Risk Assessment software
MOE	margin of exposure, i.e. the ratio of a toxicological reference dose (i.e. NOAEL) to the estimated exposure
MOET	total margin of exposure resulting from multiple chemicals and food commodities
MRL	maximum residue level
MW	molecular weight
NE	normalised exposure, i.e. the ratio the estimated exposure to a toxicological reference dose (i.e. NOAEL)
NET	total normalised exposure resulting from multiple chemicals and food commodities
NOAEL	no observed adverse effect level
PF	processing factor
PPR Panel	EFSA Panel on Plant Protection Products and their Residues

PRIMo   Pesticide Residues Intake Model
RD   residue definition
RIVM   Dutch National Institute for Health and Environment
RPC   raw primary commodity, i.e. a single-component food which is unprocessed or whose nature has not been changed by processing (e.g. apples)
RPCD   raw primary commodity derivative, i.e. a single-component food which has been physically changed by processing (e.g. apple juice)
SC PAFF Standing Committee on Plants, Animals, Food and Feed
SSD   Standard Sample Description, i.e. a harmonised data model developed by EFSA for describing analytical measurements in food and feed samples
SVF   stochastic variability factor, i.e. the ratio between a specific unit concentration drawn at random from a distribution of unit concentrations and the mean of that same distribution
VF   variability factor, i.e. the ratio between the 97.5th percentile and mean of the distribution of unit concentrations
WVF   weighted variability factor, i.e. a weighted average of the stochastic variability factors drawn for each unit within a consumed portion

Appendix A – Procedure for the allocation of active substances to the measurements

1) Select distinct combinations of raw primary commodity (RPC) and residue definition reported in the occurrence data set.
2) Identify the possible combinations of RPC, residue definition and active substance (AS) (by joining the information of the residue definitions table). Retain information on the molecular weight (MW) conversion factor, on whether this combination is exclusive or not, and on the proportion for the non-exclusive combinations.
3) Add the relevant no observed adverse effect level (NOAEL) to each combination (join information from the active substance table using the active substance as the key).
4) Identify the authorisation status for each combination (join information from the authorisations table using the RPC and active substance as the keys).

Tier I

5) There may now be combinations of RPC, residue definition and AS which refer to the same RPC and residue definition. Data are sorted by RPC, residue definition and NOAEL (ascending) and for each combination of RPC and residue definition, the first combination of RPC, residue definition and AS is retained, i.e. the one with the lowest NOAEL (most toxic AS).
6) For each measurement in the occurrence data set, the AS is assigned on the basis of the combinations derived at step 5 (using the RPC and the residue definition as keys).

Tier II

5) There may now be combinations of RPC, residue definition and AS which refer to the same RPC and residue definition. For each RPC and residue definition, only the combinations with authorised uses are retained. If none are authorised, all combinations are retained.
6) For each measurement in the occurrence data set, the AS is assigned on the basis of the combinations derived at step 5 (using the RPC and the residue definition as keys). If for a given measurement more than one AS could be assigned, only one AS is selected randomly using equal probability (regardless whether the AS is part of the cumulative assessment group).
7) For each measurement it is verified whether the combination RPC, residue definition and AS assigned is exclusive or not. If it is not exclusive:

   a) The residue value and the limit of quantification (LOQ) value are multiplied by the proportion specified in the residue definition table.
   b) The exclusive AS of that residue definition is identified (from the residue definitions table).
   c) A new measurement is generated for the same sample but for the exclusive AS identified above. The residue value and the LOQ value are also multiplied by a factor equal to (1 − proportion of the non-exclusive substance).
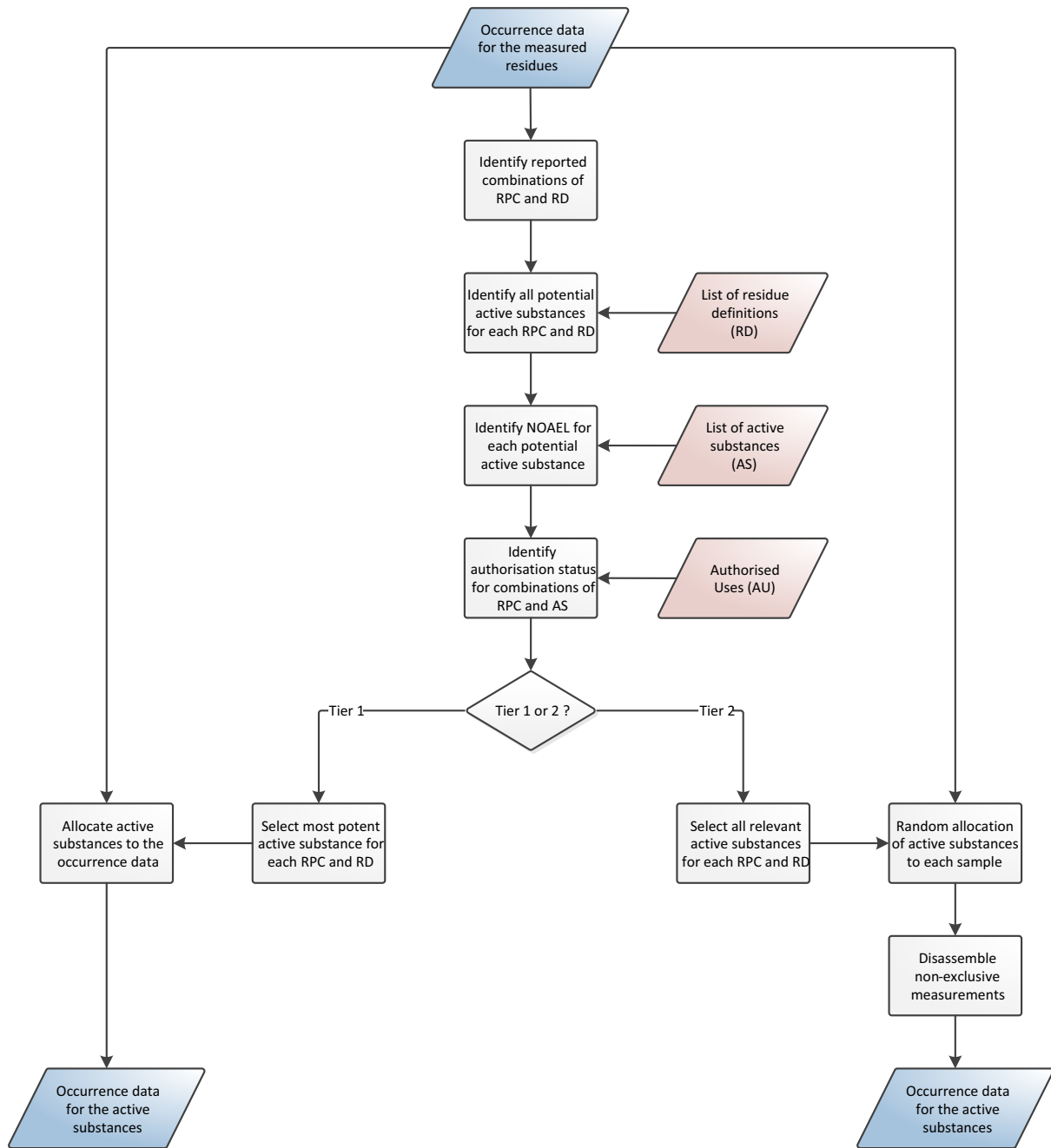
Figure A.1: Flow chart for the allocation of active substances to the measurements

Appendix B – Procedure for the extrapolation of measurements

1) Within the occurrence data set, count the number of observations per combination of active substance (AS) and raw primary commodity (RPC).

2) From the RPCs table and the ASs table, create a full matrix with all possible combinations of AS and RPC. Join this table with table created at step 1 to derive a complete list of missing and observed combination (using RPC and AS as keys).

3) Identify the maximum residue level (MRL) and the authorisation status of each of these combinations (using RPC and AS as keys). Since MRLs are defined at residue definition level, a preliminary step joins MRL table and residue definition table to associate MRL information to the active substances.

4) Identify for each combination all valid extrapolations on the basis of the extrapolation rule table. Extrapolations for a given AS and RPC are considered valid only when:

   a) the number of observations for the FromFood is equal or above 10.
   b) the number of observations for the ToFood is below 10.
   c) MRL for FromFood and ToFood is equal.
   d) Both FromFood and ToFood are authorised.

5) For each AS and for each valid extrapolation, the measurements in the FromFood are listed (can be positive or left-censored).

6) For each AS and for each valid extrapolation, the samples of the ToFood that were not analysed for the AS are listed (i.e. the missing values). This implies indeed that no extrapolation will be done if there are no samples at all for a given food.

7) Random measurements (identified at step 7) are combined with random samples (identified at step 6). This is repeated until all the FromFood measurements or all the ToFood samples are assigned. Hence, if there are insufficient measurements in the FromFood, missing values in the ToFood will remain. If there are insufficient samples in the ToFood, some measurements in the FromFood will not be assigned.

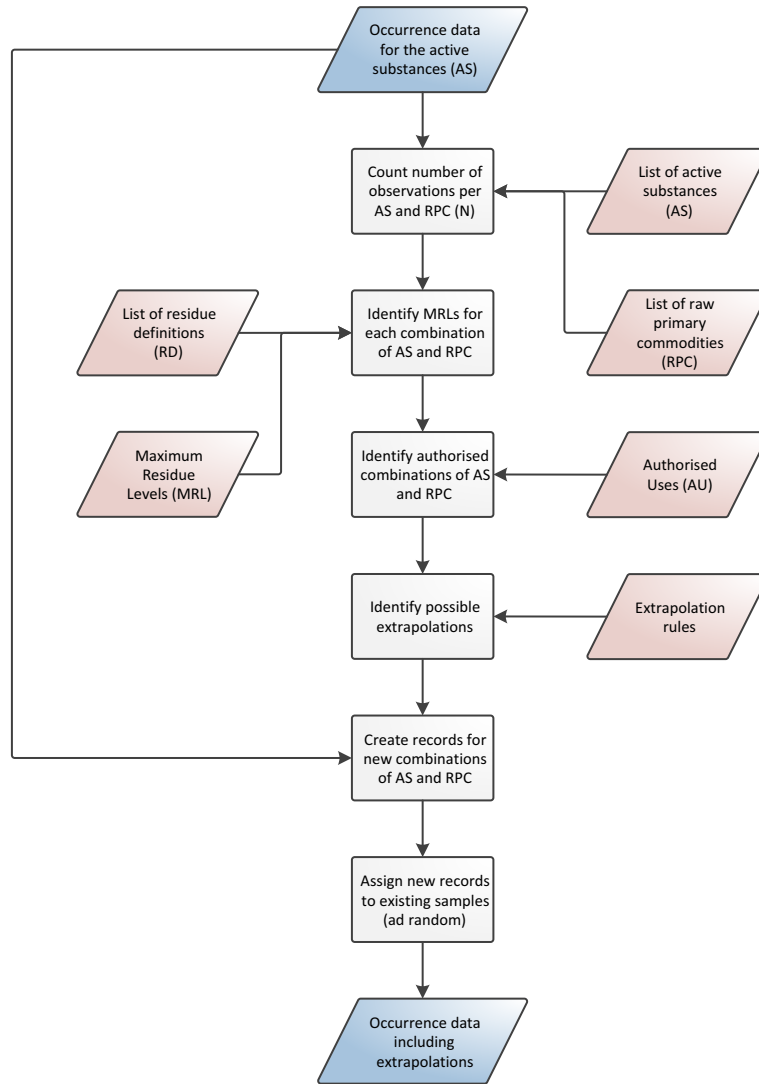8) Newly extrapolated values are added to the occurrence data set.

Figure B.1: Flow chart for extrapolation of the occurrence data

Appendix C – Procedure for the imputation of left-censored measurements

Tier I

1) Retrieve from the occurrence data set all records which refer to a quantifiable result and identify distinct combinations of raw primary commodity (RPC) and active substance (AS). This results in a list of RPC/AS combinations where the non-quantifiable results will be assumed to be at 1/2 limit of quantification (LOQ).

2) Identify in the occurrence data set all left-censored records that refer to any of the combinations listed at step 1 (using RPC and AS as keys). Assign 1/2 LOQ as a result for those records.

3) Assign zero to all remaining left-censored records in the occurrence data set.

Tier II

1) Define the list of agricultural use patterns (AUPs) observed in the data set. An AUP is the combination of AS quantified within a RPC. The list is derived as follows:

   a) Retrieve from the occurrence data set all samples which have at least one quantifiable result.

   b) Identify for each of the previous samples the AUP by concatenating the ASs quantified in each sample.

   c) Select all the distinct AUPs and assign a identifier to each AUP.

   Example: Among all apple samples, substances X, Y and Z were measured, and the following combinations were quantified within single samples: (X), (X-Y-Z), (Y), (X-Y) and (Y-Z). These combinations are now identified as AUP1, AUP2, AUP3, AUP4 and AUP5, respectively.

2) Count the number of samples for each AUP, i.e. the number of times that the AUP appears in the data set.

   Example: Number of apple samples where AUP1 was observed is 200; number of apple samples where AUP2 was observed is 23, etc.

3) Identify the analytical scope of each sample and, for each AUP, identify the number of samples where the AUP is covered by the analytical scope:

   a) From the occurrence data set, identify for each sample the analytical scope by concatenating the ASs measured in each sample.

   Example: Samples were measured either for substance Y only (Scope1), for substances X and Y (Scope2), for substances X, Y and Z (Scope3) or for substances Y and Z (Scope4).

   b) Count the number of samples for each analytical scope.

   Example: Number of samples where Scope1 was measured is 500; number of samples where Scope2 was measured is 250; number of samples where Scope3 was measured is 1,250; Number of samples where Scope4 was measured is 2,000.

   c) For each AUP, identify the analytical scopes that include all ASs of that AUP.

   Example: AUP1 is covered by Scope2 and Scope3 only.

   d) For each AUP, sum the number of samples for all analytical scopes identified at step 3c.

   Example: The number of samples where Scope2 and Scope3 were measured is 250 and 1,250. Hence, the total number of samples where AUP1 is covered by the analytical scope is 1,500.

4) Calculate frequency for each AUP (N samples AUP/N samples analytical scope).

   Example: Number of apple samples where AUP1 was observed is 200 (calculated at step 2). Number of apple samples where AUP1 is covered by the analytical scope is 1,500 (calculated at step 3). Hence, the frequency of AUP1 in apples is 13.3%.

5) Adjust frequencies for authorised AUPs (i.e. when all substances in the AUP are authorised) to obtain a total AUP frequency of 100% per RPC. This assumes that each sample in the occurrence data set was treated according to one AUP.

Example: 5 AUPs were observed in apples and frequencies for each AUP were calculated: AUP1 (13.3%), AUP2 (2.3%), AUP3 (9.8%), AUP4 (1.2%) and AUP5 (0.2%). However, only AUP1, AUP3 and AUP4 include substances that are all authorised. Therefore, only these AUPs are adjusted to obtain a total number AUP frequency of 100%. Frequencies of AUP2 and AUP5 remain unchanged and the following adjusted frequencies are obtained: AUP1 (53.4%), AUP2 (2.3%), AUP3 (39.3%), AUP4 (4.8%) and AUP5 (0.2%).

6) Calculate use frequency for each combination of RPC and AS and identify the corresponding number of measurements that should be set to 1/2 LOQ:

   a) For each combination of RPC and AS, calculate the use frequency by summing the AUP frequencies of all AUPs that contain the AS.

   Example: 5 AUPs were observed in apples and the following adjusted frequencies are obtained: AUP1 (53.4%), AUP2 (2.3%), AUP3 (39.3%), AUP4 (4.8%) and AUP5 (0.2%). Only AUP1, AUP2 and AUP4 include the use of substance X. Therefore, the estimated use frequency of substance X in apples is 60.5%.

   b) For each combination of RPC and AS, calculate the percentage of true zeros (i.e. 100 − use frequency calculated at step 6a)

   Example: If the estimated use frequency of is 60.5%, the expected percentage of true zeros is 39.5%.

   c) For each combination of RPC and AS, calculate the number of true zeros by multiplying the percentage of true zeros (calculated at step 6b) with number of measurements for that AS and RPC and divide by 100.

   Example: For substance X in apples, if the expected percentage of true zeros is 39.5% and the total number of measurements is 3,562, the estimated number of true zero measurements is 1,407.

   d) For each combination of AS and RPC, count the total number of measurements. Subtract from this value the number of samples that already have a measured value and the number of true zeroes calculated at step 6c. This is the number of samples that should be set to 1/2 LOQ. If a negative number is obtained, set to 0.

   Example: For substance X in apples, if the total number of measurements is 3562, the number of quantifiable measurements is 126 and the estimated number of true zero measurements is 1,407, the number of measurements to be imputed at 1/2 LOQ is 2,029.

7) From the left-censored data reported in the occurrence data set, randomly select for each RPC and AS the number of samples (as calculated above). Assign a residue value of 1/2 LOQ.

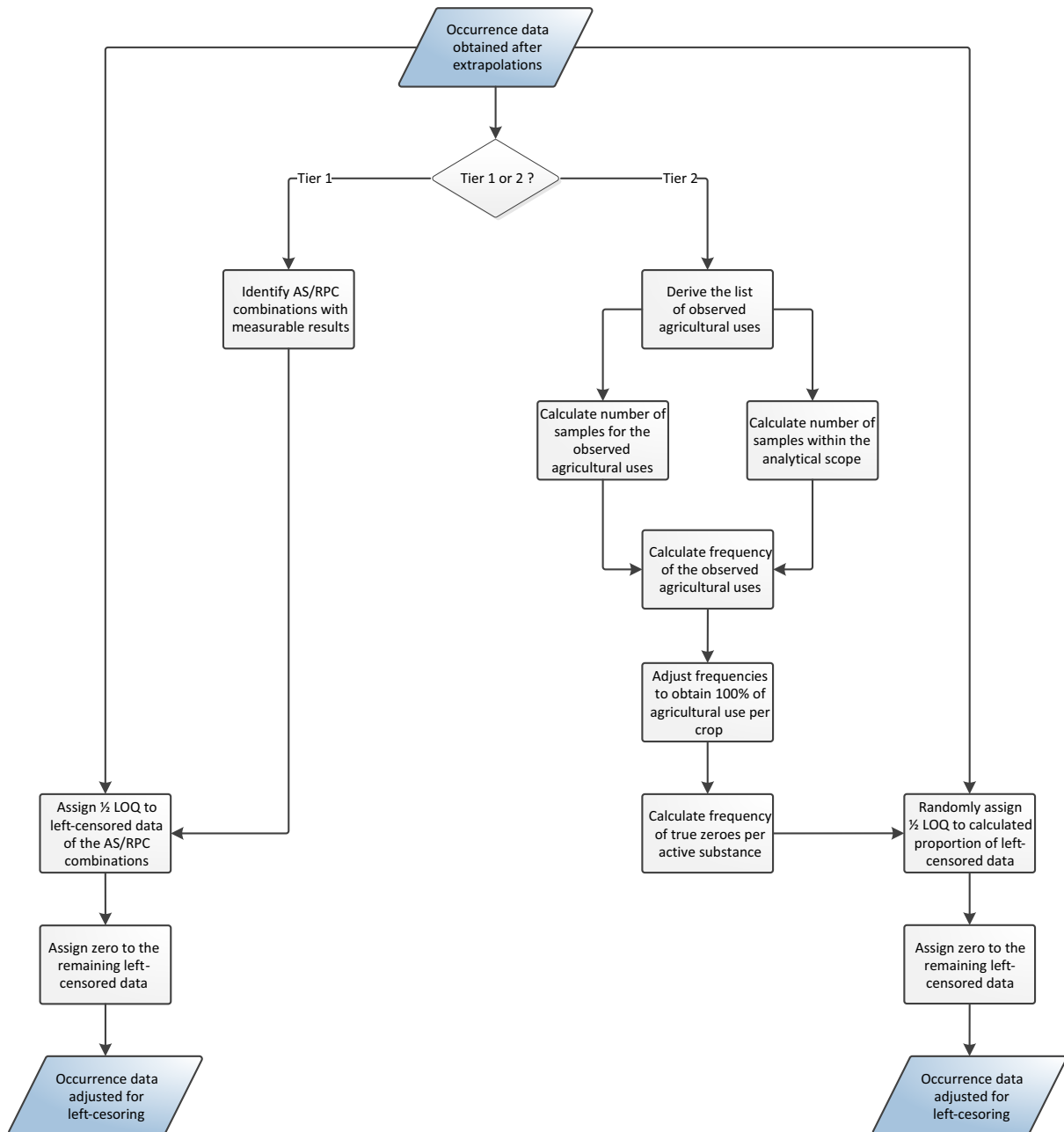8) Assign zero to all remaining left-censored records in the occurrence data set.

Figure C.1: Flow chart for the imputation of left-censored measurements

Appendix D – Procedure for the imputation of missing measurements

Tier I

1) From the occurrence data set calculate the cumulative potency (sum of concentrations of active substances divided by the respective no observed adverse effect level (NOAEL)) for each sample.

2) From the occurrence data set derive the list of samples with missing measurements for each combination of raw primary commodity (RPC) and active substance (AS). This implies that no imputation will be done if there are no measurements at all for a certain combination of RPC and AS.

3) Assign cumulative potency to samples with missing measurements created at step 2 by joining the information created at step 1 and order all measurements from high to low according to their cumulative potency for each RPC and AS.

4) Count number of missing measurements (N) for each RPC and AS.

5) Drawn N measurements at random from existing samples for each RPC and AS and order the generated measurements from high to low for each RPC and AS.

6) Create N new records by joining measurements generated at step 5 to samples identified at step 3. This implies that the highest extracted measurements will be assigned to the most potent samples.

7) Add records for missing measurements to the occurrence data set.

Tier II

1) From the occurrence data set derive the list of samples with missing measurements for each combination of RPC and AS:

   a) Create records for implicit zero measurements of ASs associated to unspecific residue definitions that were considered not present in a sample and add these records to the occurrence data set.

   b) From the occurrence data set created at step 1a derive the list of samples with missing measurements for each combination of RPC and AS. This implies that no imputation will be done if there are no measurements at all for a certain combination of RPC and AS.

2) Count number of missing measurements (N) for each RPC and AS.

3) Drawn N measurements at random from existing samples for each RPC and AS.

4) Create N new records by randomly assigning the measurements extracted in step 3 to each sample identified at step 1b and add records for missing measurements to the occurrence data set.
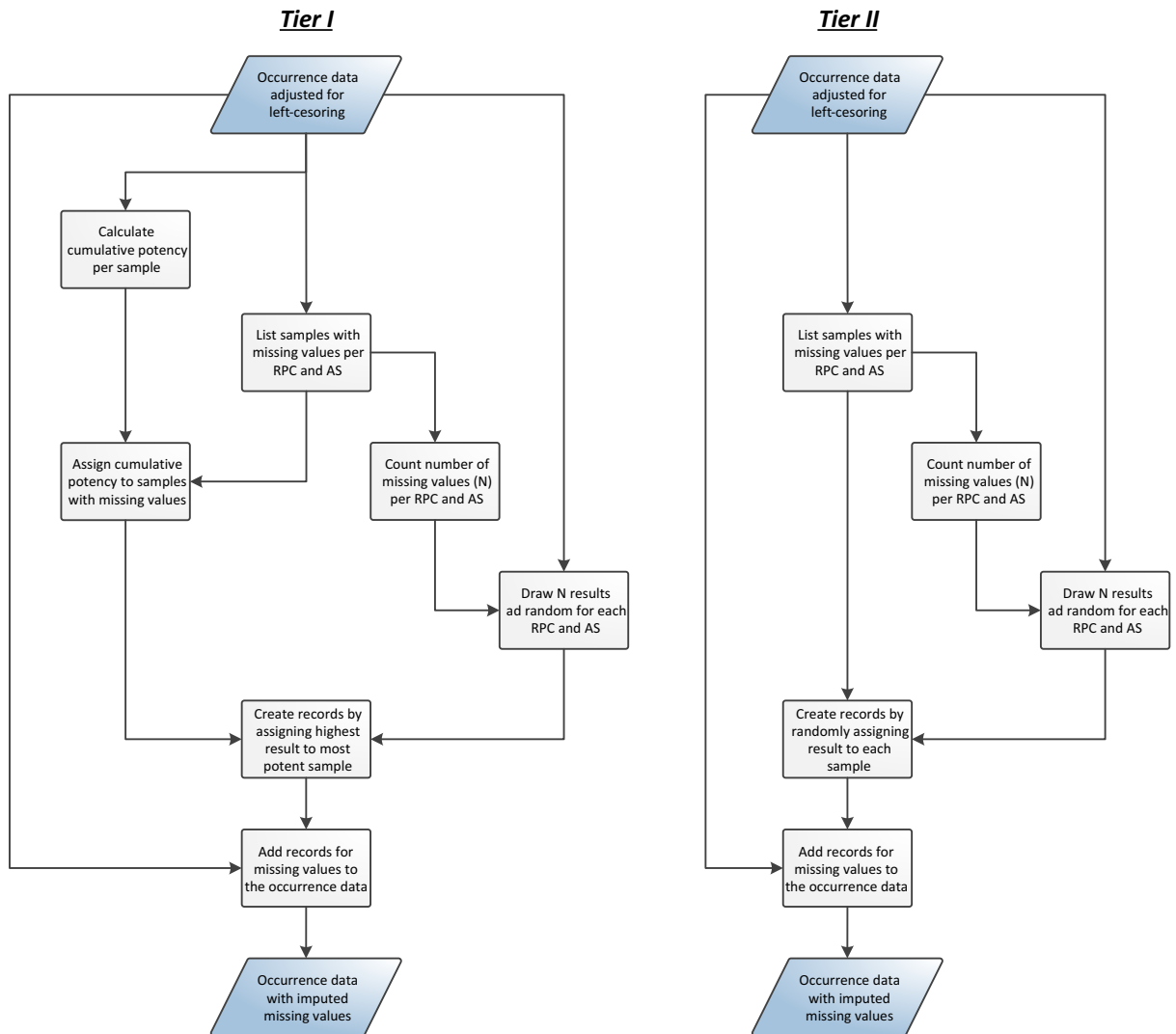
Tier I

Occurrence data adjusted for left-cesoring

Calculate cumulative potency per sample

List samples with missing values per RPC and AS

Assign cumulative potency to samples with missing values

Count number of missing values (N) per RPC and AS

Draw N results ad random for each RPC and AS

Create records by assigning highest result to most potent sample

Add records for missing values to the occurrence data

Occurrence data with imputed missing values

Tier II

Occurrence data adjusted for left-cesoring

List samples with missing values per RPC and AS

Count number of missing values (N) per RPC and AS

Draw N results ad random for each RPC and AS

Create records by randomly assigning result to each sample

Add records for missing values to the occurrence data

Occurrence data with imputed missing values

Figure D.1: Flow chart for the imputation of missing measurements

Appendix E – Procedure for deriving the acute exposure distribution

1) From the consumption data set, randomly select (with replacement) 100,000 subject days for each dietary survey. For each resampled subject day, all consumption events within the day are retained.

2) For each raw primary commodity (RPC) consumed within the individual subject days, assign random samples extracted from the occurrence data set (with replacement).

3) Assign processing factors (PFs) to the relevant records of the data set created at step 2 by joining information from the PF table (using the RPC, active substance and FoodEx2 facet as the keys). If no PF is available for a specific combination, then a missing value is assigned to the PF.

4) Incorporate variability factor (VF) to the relevant records of the data set created at step 3:

   a) Calculate the two parameters a and b of a beta distribution defined over the interval $(0, nUnits)$, where nUnits is the number of units per sample as defined in Section 2.2.1. An iterative method is applied to find a such that the cumulative probability $P[Beta(a, b)] = 0.975$ for $b = a(nUnits - 1)$.

   Tier I

   A VF of 5 or 7 is used to derive the beta distribution.

   Tier II

   A VF of 3.6 is used to derive the beta distribution.

   b) Join information from the processing types table (using FoodEx2 facet as the key), the RPC table (using RPC as the key) and the parameters a and b derived at step 4a to the data set created at step 3.

   c) For each active substance (AS) measured in the sample assigned to the consumption event randomly derive a stochastic VF (SVF) from the beta distribution defined at step 4a. When the consumed portion is composed of multiple units, multiple SVFs are drawn from the same beta distribution in order to calculate a weighted VF (WVF) as described in Section 2.4.2.

   d) An adjusted concentration is calculated by multiplying the concentration measured in the sample with the WVF (or SVF) calculated at step 4c. If no VF is available for a certain RPC or if the processing type involves bulking/blending, then the adjusted concentration is equal to the concentration measured in the sample.

5) Calculate normalised exposure (NE) for each record using formula described in Section 2.4.2 to obtain NE per subject day, RPC and AS.

6) Sum all normalised exposures of RPCs and AS per subject day to obtain a total normalised exposure (NET) for each subject day.
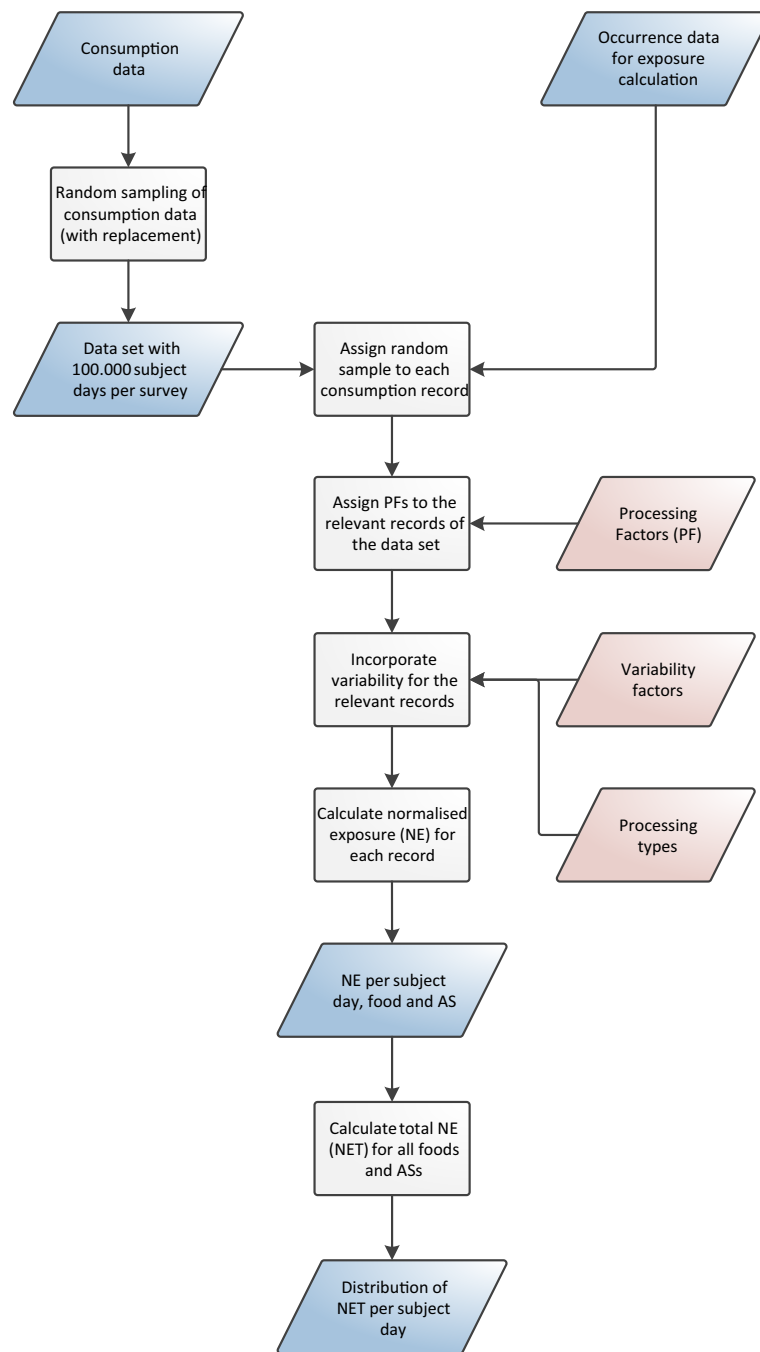
Figure E.1: Flow chart for the calculation of acute exposure

Annex A.1 – Input data for the exposure assessment of CAG-NAN

Annex A.1 can be found online on EFSA's knowledge junction: https://doi.org/10.5281/zenodo.3338270

Annex A.2 – Input data for the exposure assessment of CAG-NAM

Annex A.2 can be found online on EFSA's knowledge junction: https://doi.org/10.5281/zenodo.3338270

Annex B.1 – Output data from the Tier I exposure assessment of CAG-NAN

Annex B.1 can be found online on EFSA's knowledge junction: https://doi.org/10.5281/zenodo.3338270

Annex B.2 – Output data from the Tier I exposure assessment of CAG-NAM

Annex B.2 can be found online on EFSA's knowledge junction: https://doi.org/10.5281/zenodo.3338270

Annex C.1 – Output data from the Tier II exposure assessment of CAG-NAN

Annex C.1 can be found online on EFSA's knowledge junction: https://doi.org/10.5281/zenodo.3338270

Annex C.1 – Output data from the Tier II exposure assessment of CAG-NAN

Annex C.2 – Output data from the Tier II exposure assessment of CAG-NAM

Annex C.2 can be found online on EFSA's knowledge junction: https://doi.org/10.5281/zenodo.3338270