MCRA 7

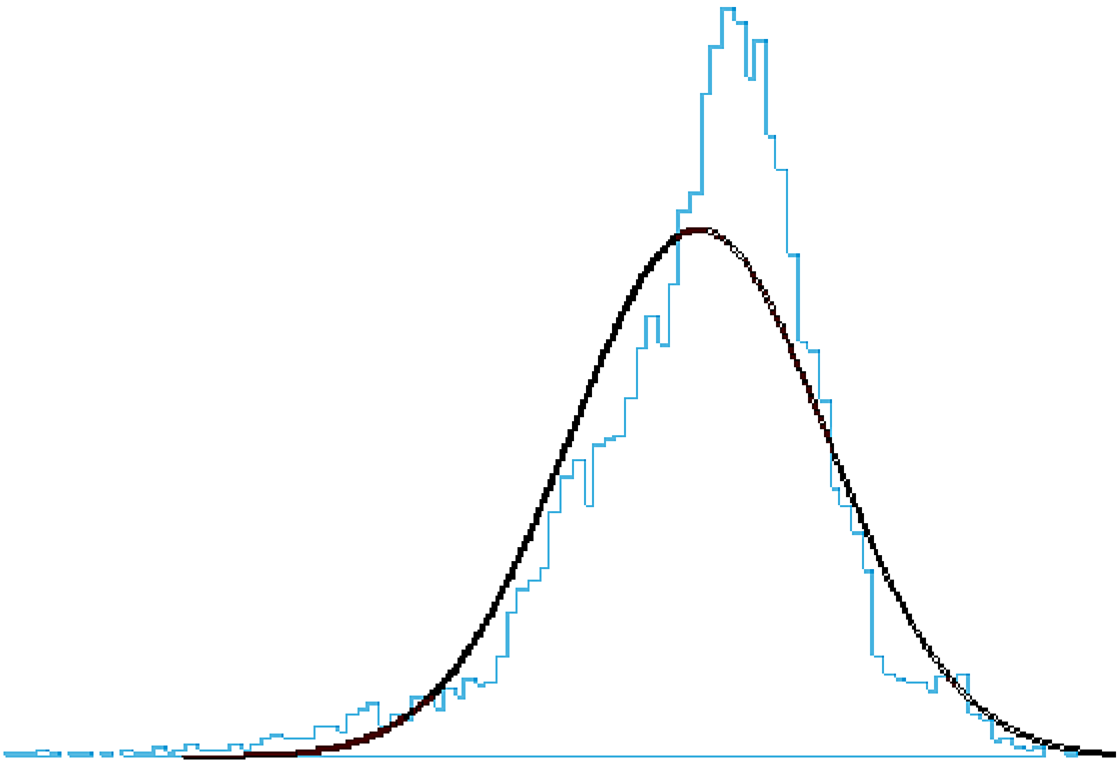
a web-based program for Monte Carlo Risk Assessment

Data Formats 2011-12-19

documenting MCRA Release 7.1

Waldo J. de Boer
Hilko van der Voet

Contributions to MCRA were made by: Paul W. Goedhart, Gerie W.A.M. van der Heijden, Sanne Korzec, Helen Owen (Food and Environment Research Agency, UK), Jac T.N.M. Thissen



Biometris
Wageningen University and Research centre
RIVM
Centre for Substances and Integrated Risk Assessment,
National Institute for Public Health and the Environment

19 December 2011

Biometris is the unit for Mathematical and Statistical Methods of Wageningen University & Research centre.

Post address
P.O. Box 100
6700 AC Wageningen
The Netherlands
Visiting address
Bornsesteeg 47, building no. 116
6708 PD Wageningen
Telephone: +31 (0)317 476925
Telefax: +31 (0)317 483554

RIVM National Institute for Public Health and the Environment

Post address
P.O. Box 1
3729 BA Bilthoven
The Netherlands
Visiting address
Antonie van Leeuwenhoeklaan 9
3721 MA Bilthoven
Telephone: 0031 30 2749111
Telefax: 0031 30 2742971

1 Preparing the data

1.1 Overview

MCRA requires that all information needed for a risk assessment is stored in MS Access databases or Excel. The database approach requires some understanding about the relational structure of the data, in exchange flexibility to pre-process the data is offered and the results of the analysis have greater detail. The Excel type of data have a simple lay-out, recording the intakes of individuals together with characteristics on *e.g.* age and or sex.

MS Access data are organised into tables which are divided into 3 groups with information on

- food consumption data,
- concentration data,
- linking and additional data.

See Table 1 for an overview. To run MCRA, tables called 'required' should always be supplied. Selection of 'additional' tables depends on subset requirements and model specifications.

Required tables (minimal configuration)	
Food	food codes and labels (see 1.5.1)
Food consumption data	
FoodConsumption	consumption of foods, foods as consumed (see 1.6.1)
Individual	individual characteristics (see 1.6.2)
Concentration data	
Compound	substance codes, labels, agricultural and toxicological limits (see 1.7.1)
Country	country codes, labels (see 1.7.2)
ConcentrationValues	concentration data (full data) (see 1.7.3)
Additional tables (for querying or specific options)	
FoodProperties	food codes and labels, and food specific information(see 1.5.2)
FoodComposition	food codes and labels, compositions (see 1.5.3)
FoodMarketShare	food codes and labels, marketshares (see 1.5.4)
FoodConsumptionQuantification	units and quantification (see 1.5.5)
FoodConsumptionUncertainty	units and uncertainty (see 1.5.6)
Food consumption data	
FoodSurvey	Name of survey (see 1.6.3)
ProcessingType	processing codes and labels (see 1.6.4)
Processing	processing factors (see 1.6.5)
Concentration data	
VariabilityProd	unit variability factors (see 1.7.4)
VariabilityCompProd	unit variability factors, substance-specific (see 1.7.5)
VariabilityProcCompProd	unit variability factors, processing- and substance-specific (see 1.7.6)
AgriculturalUse	information on the agricultural use of substancess (*e.g.* use allowed, percent crop treated) (see 1.7.7)
ConcentrationWorstcaseValues	information on worstcase values (*e.g.* substance and food specific worstcase values) (see 1.7.8)

Table 1: Overview of tables

Tables are organised into columns (fields) and rows (records). In the next paragraphs, the format of tables is described.

General remarks:

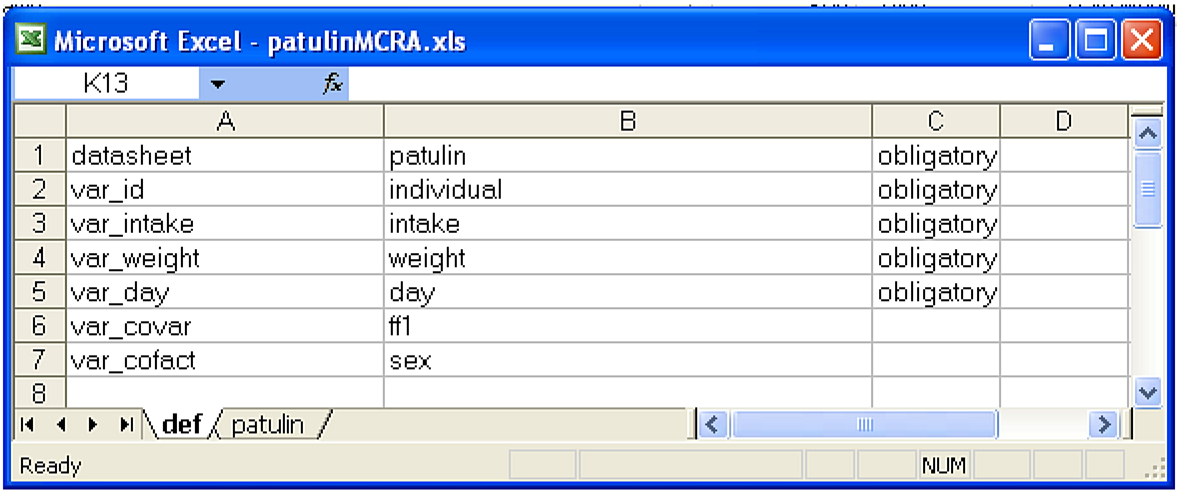
- Table and column names should be exactly as indicated in the sections below and each table should contain **all** fields, unless otherwise stated (optional).
- Missing values are indicated with code 9999, unless otherwise stated. In general, an empty cell is also interpreted as a missing value. Occasionally, the use of empty cells leads to errors in retrieving data. Therefore, it is advised to use the code 9999 to indicate missing values

In sections 1.5 , 1.6 and 1.7 the format is explained: the table name is given followed by field names and a description with in parentheses the datatype. Each section ends with some notes and an example.

1.2 Excel

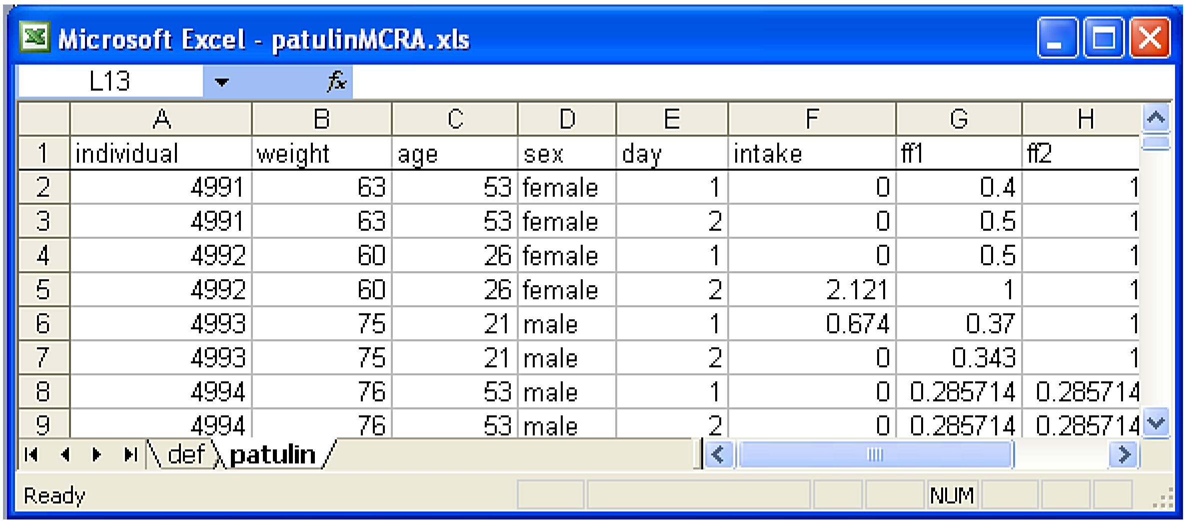
When data are organised on a per individual basis, *e.g*. the total intake for an individual *i* on day *j* is recorded, Excel spreadsheets may be used as input to MCRA. The Excel file contains two worksheets, one defining the layout of the data, one containing the data.
Rules are:

- name of the definition sheet is a system name: **def** (do not change)
- cell A1 to A7 are system names (do not change)
- cell A1 to A5 are obligatory
- var_covar and var_cofact are optional
- cell B1 to B7 are user defined names
- cell B1 defines the datasheet (here patulin)
- column names in the datasheet are defined in the definition file
- all intakes including zero intakes are reported
- all individuals have the same number of days
- datasheet may contain redundant columns



	A	B	C	D
1	datasheet	patulin	obligatory	
2	var_id	individual	obligatory	
3	var_intake	intake	obligatory	
4	var_weight	weight	obligatory	
5	var_day	day	obligatory	
6	var_covar	ff1		
7	var_cofact	sex		
8				

Figure 1: definition sheet: def



	A	B	C	D	E	F	G	H
1	individual	weight	age	sex	day	intake	ff1	ff2
2	4991	63	53	female	1	0	0.4	1
3	4991	63	53	female	2	0	0.5	1
4	4992	60	26	female	1	0	0.5	1
5	4992	60	26	female	2	2.121	1	1
6	4993	75	21	male	1	0.674	0.37	1
7	4993	75	21	male	2	0	0.343	1
8	4994	76	53	male	1	0	0.285714	0.285714
9	4994	76	53	male	2	0	0.285714	0.285714

Figure 2: data sheet: patulin

1.3 Harmonised CODEX codes

In the MCRA program we use harmonised CODEX codes in the interest of Pan-European risk assessment. This coding offers flexibility to enter food consumption data and concentration data at any desired level of food coding (*e.g.* food as consumed, ingredient, foods as measured, raw agricultural commodity, processed food, brand level, etc.).

1.3.1 Definitions

A food code is a string consisting of symbols:

- letters (case-unsensitive, so x and X are the same letter),
- digits, and/or
- special symbols, such as ~!#$^&*()+-=[]{};’:”,./<>?`

Some special symbols are reserved for special use (see below), and can not be used freely in own codes:

- &
- $
- -
- *

Some symbols are not allowed at all, because this would interfere with the way the strings are analysed:

- %
- _
- @

The first symbol should be:

- a letter (indicates a CODEX code or a code derived from a CODEX code), or
- & followed by a 2-letter country code (indicates a national food code)

CODEX codes start with two letters and four digits, and should comply with the CODEX Classification of foods and animal feeds. The code XX9999 (usually followed by a subtype code) can be used for all foods which cannot be placed in the Codex classification system.

Any code (CODEX code or national code) can be followed by:

- $ plus a subtype code, and/or

- - plus a processing code
- *- plus a processing code. Here the asterisk (*) serves as a wildcard for the preceding code: the processing information is valid for all codes that start with the code preceding the *.

Subtype codes and processing codes can have any format. Multiple levels of subtype code are allowed (*e.g.* &NL00$123$456). Only one level of processing code is allowed (*e.g.* FP0226-2). Subtype codes should precede processing codes (*e.g.* &NL00$123$456-2).

Within EU-Safefoods we will harmonise subtype codes and processing codes as far as they apply to CODEX codes. For this purpose lists will be maintained at RIKILT, and any new subtype code or processing code is checked against this list, and when found appropriate added to this list.

The table Food contains descriptive names for <u>all</u> food codes that are used in the tables **FoodConsumption** and **ConcentrationValues** and that are to be to included in the analysis. Names will be in English, but an additional column with alternative names can be used.

1.3.2 Conversion of codes

We distinguish 3 types of food code conversion, which may be provided in three different tables:

1. Food processing (table **Processing**)
   Processing factors will be applied to concentration data.
2. Food composition (table **FoodComposition**)
   Composition percentages will be used to transform the consumed amounts.
3. Subtypes/ Market share (table **FoodMarketShare**)
   Market share percentages will be used as probabilities for selecting concentration data for each of the subtypes.

1.3.3 Conversion rules

- For each *code* in table **FoodConsumption,** , try to find the most appropriate concentration information by searching information in a specific order according to the steps in the scheme below.
- If a code is converted into one or more other codes, then for any such other code re-start the search scheme before continuing to the next step or substep in the scheme with the old code.
- The search is ended if concentration data have been found for *code* or as many as possible derived codes. When a code is converted to multiple new codes (composition, subtypes), then the search continues for all these new codes.
- If no link can be made to concentration data, then the consumption of this *code* is considered irrelevant for the current intake assessment.

In 1.3.4 a scheme is given to link the food consumption and concentration data.

1.3.4 Scheme to link food consumption and concentration data

Find in Figure 3, an schematic outline of the search for food codes. After a successful search, the code is found in the ConcentrationValues table. If the code is not found, searching starts in one of the additional/link tables. If the code is found, the search starts again in the ConcentrationValues table and the search is repeated. If a code is not found at all, the search for a code is unsuccessful.

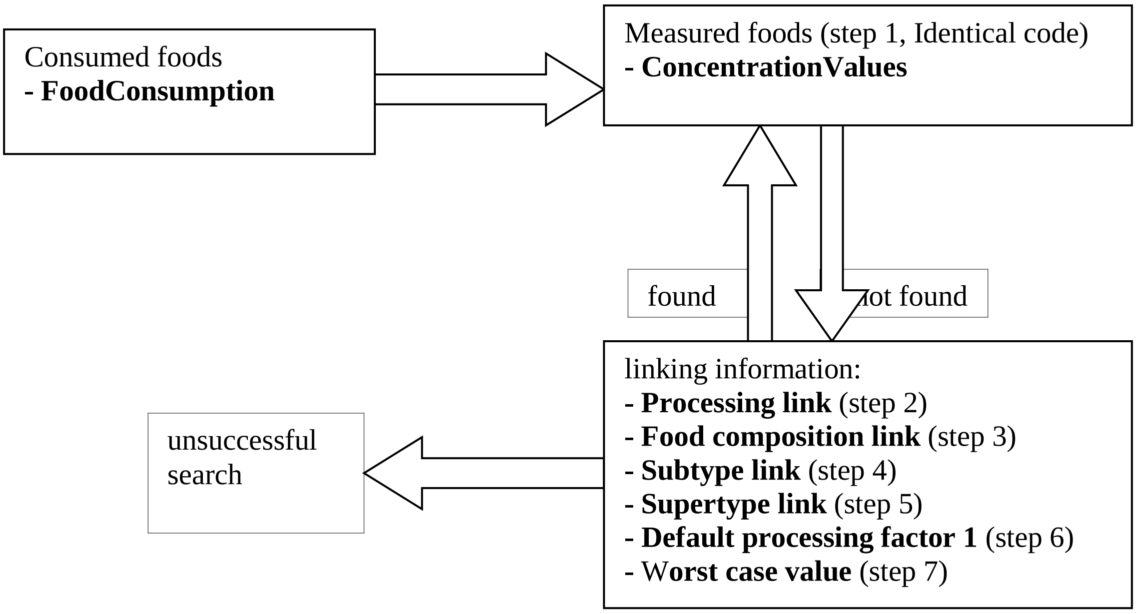


Figure 3: Link scheme to find CODEX codes

Step 1. (Identical code)
Try to find *code* in the field **foodmeasured** of the **ConcentrationValues** (1.7.3) table.
If found, the search is ended successfully.

Step 2. (Processing link)

a. Try to find *code* in the field **foodprocessed** of the **Processing** table (1.6.5), and convert to the code specified in the field **foodunprocessed**.
b. (wildcard match) Try to find a wildcard match code in the field **foodprocessed** of the **Processing** table. Wildcard match codes consist of an initial string (*startcode*, may be empty), an asterisk (*), and possibly a processing part (-*processingtype*). * may be any string *endcode* (not containing a -) such that *code* equals *startcodeendcode* or *startcodeendcode-processingtype*.
    a. If *code* contains a processing part (-*processingtype*), then the wildcard match code should also end with -*processingtype.* Convert to the code specified in the field **foodunprocessed**, where *endcode* is substituted for any * in the new code.
    b. If *code* contains no processing part, then the wildcard match code should also contain no processing part. Convert to the code specified in the field **foodunprocessed**, where *endcode* is substituted for any * in the new code.
c. A check is made on the simultaneous use of Food conversion factors for weight changes between two foods, e.g. a food as eaten and a food as measured (step 3 of the conversion algorithm) and processing factors for concentration changes between these two foods (step 2 of the conversion algorithm). If the same code pair is found in both tables, the processing factor is modified to indicate the change in amount (rather than concentration) of the compound.

Step 3. (Food composition link)

a. Try to find *code* in the field **food** of the **Foodcomposition** table (1.5.3), and convert to one or more ingredient codes found in the field **ingredient**
b. If *code* contains a processing part (*maincode-processingtype*), then try to find *maincode* in the field **food** of the **Foodcomposition** table. Convert to one or more ingredient codes and add *-processingtype* to the new codes.

Step 4. (Subtype link)
Starting from *code*, try to find subtype codes equal to *code$** in the field **foodtype** of the **FoodMarketShare** table (1.5.4 , where the strings represented by the wildcard are not allowed to contain a $ themselves (meaning that we look only for codes one level down in the type hierarchy). Check that for the selected codes the market share percentages in the field **marketshare%** sum to 100%.

If this is not OK, then the result depends on the user option "Allow conversion to subtypes not summing to 100% and rescale".

1. If this option is allowed, then the marketshare percentages are rescaled to a sum of 100 %.
2. If this option is not allowed, then the search in step 4 is considered unsuccessful.

Step 5. (Supertype link)
This step will only be taken if the user has explicitly allowed this option.

a. If *code* contains a subtype part but no processing part (*maincode$subcode*), then convert to *maincode*.
b. If *code* contains a subtype part and a processing part (*maincode$subcode-processingtype*), then convert to *maincode-processingtype*.

Step 6. (Default processing factor 1)
If *code* contains a processing part (*-processingtype*), then remove it.

Step 7. (Worst case value)
This step will only be taken if the user has explicitly allowed this option.
Try to find *code* in the field **food** of the **ConcentrationWorstcaseValues** table (1.7.8). If found, the search is ended successfully.

1.3.5 Example of use

The **FoodConsumption** table (1.6.1) may contain &NL070251 (Apple pie Dutch):

individual	dayofsurvey	foodconsumed	amountconsumed	foodsurvey
1012	1	&NL070251	150	DNFCS

If measurements are available for FP0226$Elstar (Apple Elstar), FP0226$JonaGold (Apple JonaGold) and GC0654 (Wheat), then we need a conversion.

The **FoodComposition** table (1.5.3) may then specify the composition data that apple pie contains peeled apple and wheat:

food	ingredient	proportion%
&NL070251	FP0226-2	58.09
&NL070251	GC0654	14.52

The **Processing** table (1.6.5) may contain a processing factor for peeling of apples:

compound	foodprocessed	foodunprocessed	proctype	procnom	procupp

120701	FP0226-2	FP0226	2	0.3	0.5

where the field **proctype** specifies explicitly the type of processing (2 = peeling in this case), and fields **procnom** and **procupp** are processing factor nominal and upper values.
The **FoodMarketShare** table (1.5.4) may contain market shares for subtypes of apple:

foodtype	marketshare%
FP0226$Elstar	30
FP0226$JonaGold	70

Finally, the **ConcentrationValues** table (1.7.3) should then contain measured concentrations for FP0226$Elstar (Apple Elstar), FP0226$JonaGold (Apple JonaGold) and GC0654 (Wheat):

compound	foodmeasured	year	month	samplingtype	country	numberofsamples	value
120701	FP0226$Elstar	2006	11	M	NL	1	0.34
120701	FP0226$Elstar	2006	11	M	NL	6	-0.01
120701	FP0226$JonaGold	2006	11	M	NL	1	0.20
120701	FP0226$JonaGold	2006	11	M	NL	1	0.05
120701	GC0654	2006	11	M	NL	6	-0.01

The **Food** table (1.5.1) should at least contain the following entries:

food	foodname
&NL070251	Apple pie Dutch
FP0226$Elstar	Apple Elstar
FP0226$JonaGold	Apple JonaGold
GC0654	Wheat

Note that entries for intermediate stages such as FP0226-2 or FP0226 are not obligatory.

1.4 Portion size estimation

Standardization of methods may increase the reproducibility within and between surveys, but recorded dietary intake data are still subject to measurement error due to portion size estimation. Therefore, quantification of this uncertainty is useful and contributes to the interpretation of the outcomes and conclusions.

1.4.1 Definitions

The statistical model of consumption portion uncertainty is based on quantification methods as used in the EPIC-SOFT program. The modelled sources of uncertainty are:

1) Uncertainty in unit weights (for EPIC-SOFT quantification methods P, H, U, S and ?);
2) Uncertainty in consumed amounts (for EPIC-SOFT quantification methods G, P, H and U)

The following scheme summarizes the new coding of the proposed database design.

Quant. method	**unit**		*example*	**amount**		*example*
	Asked scale	status of levels		Asked scale	status of levels	
G	not	not		continuous	estimate	~120 g
P	ordered series	exactly known	~ photo 2 (135 g)	continuous	estimate	~1.5

H	nominal	estimate	bowl (~100 g) ***(if needed convert ml to g)***	continu ous	estimate	~1.5
U	ordered series / nominal	estimate	small cucumber (~80 g)	continu ous	estimate	~1.5
S	not	estimate	grated cheese on pasta (~20 g)	not	not	
?	not	estimate	Salad dressing (~10 g)	not	not	

Sometimes 'unit' is sufficient to identify the food as consumed, and therefore, only this code is needed to specify the unitweight (in grams). For method P this is true, e.g. unit P4-3 is sufficient to identify the food with corresponding unitweight 125g. For method H, this is not always the case, e.g. unit H034 identifies a small bowl but the corresponding unitweight in grams may be dependent on the food. E.g. a small bowl with applesauce (fc4) has a different unitweight (here 80 g) than a bowl filled with lettuce (fc2), i.c. 35 g. Therefore, table '**FoodConsumptionQuantification**' contains a column 'food' that is needed for those units where the unitweights are dependent on the food. For unit codes that identify the food uniquely e.g. P4, the column 'food' is strictly speaking redundant, however the use of this column introduces flexibility that may be used in special cases (for example, if photos of cauliflower were used in the case of broccoli).

1.4.2 Scheme to link foods and units to quantification and uncertainty measures

To assess the uncertainties in portion size, a search in table **FoodConsumptionQuantification** and **FoodConsumptionUncertainty** is started to find unitweights and corresponding uncertainties (%) in unitweights and, if relevant, amounts. Note that each label specified in column 'unit' in table **'Foodconsumption'** should have a record in table **'FoodConsumptionQuantification'** (the reverse is not true). The label 'general' specifies a default value for the unitweight or uncertainty when specific information for combinations of food and unit as specified in table '**FoodConsumption'** is not available.
Unitweights are found using a two step search procedure, starting with the most specific information:

1) food = fcxx, unit = Qxxx, if not found proceed with
2) food = general, unit = Qxxx, if not found there is an error condition.

Uncertainties are found using a three step search procedure, starting with the most specific information:

1) food = fcxx, unit = Qxxx, if not found proceed with
2) food = general, unit = Qxxx, if not found proceed with
3) food = general, unit = general. if not found there is an error condition

Above scheme implies that not every combination of food and unit in table '**FoodConsumption'** needs a specification in table **'FoodConsumptionUncertainty'.**

1.4.3 Example of use

Table **'FoodConsumptionQuantification**' records the unitweights of each unit in grams. When unit sizes are specified in ml, these values have to be converted to grams externally.

food	unit	unitweight
general	P4-1	50
general	P4-2	75
general	P4-3	125
general	H034	88
general	Scode	20

fc4	H034	80
fc2	H034	35

- Record 1 is read as: the unit weight of unit P4-1 is 50 g
- Examples:
    - fc4, H034: unitweight = 80 g (found in step 1).
    - fc3, H034: unitweight = 88 g (found in step 2).

Table **'FoodConsumptionUncertainty'** specifies information on uncertainties for unitweight and amount.

food	unit	unitweight (%)	amount (%)
fc2	P4-1	10	5
fc2	P4-2	15	5
fc2	P4-3	15	10
fc2	H034	20	10
fc2	Scode	5	5
general	H034	10	15
fc2	H034	13	5
fc4	H034	10	8
general	general	5	10

- Record 1 is read as: for P4-1, the uncertainty for unitweight is 10%, the uncertainty for amount is 5%.
- Examples:
    - fc4, H034: unitweight = 10%, amount = 8% (found in step 1)
    - fc3, H034: unitweight = 10%, amount = 15% (found in step 2)
    - fc4, U1-003: unitweight = 5%, amount = 10% (found in step 3)

1.5 Food tables

1.5.1 Food (compulsory)

field name	description
food (text) foodname (text) foodname2 (text, optional)	food code food label alternative food label, *e.g.* national language

- Foodname2 is used for alternative foodnames.

1.5.2 FoodProperties (optional, (for unit variability compulsory))

field name	description
food (text) foodname (text) unitweight (number) edibleportion (number) largeportion (number)	food code food label nominal weight of a unit (gr) edible portion (corrected large portion weight, gr) weight of a large portion (gr)

- For unknown nominal unit weight use value 0.
- Missing values for edibleportion and largeportion: 9999.

1.5.3 FoodComposition (optional)

field name	description
food (text) ingredient (text) proportion% (number)	food code ingredients of the food proportion of each ingredient in the food (in percentages)

- Specifies the composition of foods and corresponding proportions.

1.5.4 FoodMarketShare (optional)

field name	description
foodtype (text) marketshare% (number)	subtype of food market share of each subtype (in percentages)

- Specifies food marketshares of subtypes.

1.5.5 FoodConsumptionQuantification (optional)

field name	description
food (text) unit (text, optional)) unitweight (number)	food code unit name unit weight/portion size (g)

- No missing values allowed
- When unit sizes are specified in ml, these values have to be converted to grams externally.
- Units may depend on food.
- The label 'general' specifies a default value for the unitweight when specific information for combinations of food and unit as specified in table '**FoodConsumption'** is not available.

1.5.6 FoodConsumptionUncertainty (optional)

field name	description
food (text) unit (text, optional)) unitweight% (number) amount% (numbert)	food code unit name uncertainty in unit weight /portion size (%) uncertainty in amount consumed (%)

- No missing values allowed.
- The label 'general' specifies a default value for the uncertainty when specific information for combinations of food and unit in table '**FoodConsumption'** is not available.

1.6 Food consumption tables

1.6.1 FoodConsumption (compulsory)

field name	description
individual (number) dayofsurvey (number) foodconsumed (text) amountconsumed (number) foodsurvey (text) unit (text, optional))	individual identification number day (sequential number in food consumption survey) food code consumed portion of food (g) name of survey unit name

- Contains data on consumed foods. Days without consumptions are not recorded. The number of available days per individual is inferred from this table and is assumed to be the same for each individual in the survey.
- No missing values allowed.

1.6.2 Individual(compulsory)

field name	description
individual (number) foodsurvey (text) age (number) weight (number) sex (text)	individual identification number name of survey age (*e.g.* in years, months or days) body weight (*e.g.* in kg or g) gender

- Specify in table FoodSurvey (see 1.6.3) the unit for age and weight.
- No missing values allowed.

1.6.3 FoodSurvey (optional)

field name	description
foodsurvey (text) year (number) country (text) agein (text) weightin (text)	name of survey year of survey country of survey unit of age unit of weight

- Defines characteristics of the survey.
- No missing values allowed.

1.6.4 ProcessingType (optional, (for processing compulsory))

field name	description
proctype (number) procname (text) disttype (number)	code of processing type description of processing type indicator (1/2): • simulated processing factors are restricted to the interval (0,1) using a logistic-normal distribution (1), • or simulated processing factors are restricted to positive values using a log-normal distribution (2)
bulkingblending (number)	indicator (0/1): for types of processing applied on large batches, *e.g.* juicing, sauce/puree (obligatory), • 0 = no bulking/blending ; • 1 = bulking/blending

- Information on bulking and blending is only relevant for modeling of processing effects in combination with unit variability and IESTI calculations, but should always be present in the table even when these effects are not explored.
- No missing values allowed.

Example

proctype	procname	disttype	bulkingblending
1	RAW	1	0
2	PEELING	1	0
3	COOKING IN WATER	1	0
4	BAKING OF BREAD	1	0
5	CANNED/CONSERVED	1	0
6	BREWING	1	0
7	DRYING	2	0
8	FRYING/BAKING IN FAT	1	0
9	JUICING	1	1
10	MILLING	2	0
11	MARMELADE/JAM	1	1
12	OIL EXTRACTION	2	0
13	SAUCE/PUREE	2	1
14	CLEANING	1	0
15	WASHING/CLEANING	1	0
16	WINE MAKING	2	0
99	UNKNOWN	1	0

1.6.5 Processing (optional)

field name	Description
compound (text) foodprocessed (text) foodunprocessed (text) proctype (number) procnom (number)	compound code food code processed food code unprocessed code of processing type nominal value (best estimate of 50^{th} percentile) of processing factor (defines median processing factor)
procupp (number, optional)	upper value (estimate of 95^{th} percentile or "worst case" estimate) of processing factor due to variability (from procnom and procupp a standard deviation for variability of the processing factor is derived)
procnomuncupp (number, optional)	upper 95^{th} percentile of procnom due to uncertainty (from procnom and procnomuncupp a standard deviation for uncertainty of procnom is derived)
procuppuncdf (number, optional)	degrees of freedom of a chi-square distribution describing the uncertainty of the standard deviation for variability of the processing factor
procuppuncupp (number, optional)	upper 95^{th} percentile of procupp due to uncertainty (from procnom, procupp, procnomuncupp and procuppuncupp the degrees of freedom of a chi-square distribution describing the uncertainty of the standard deviation for variability is derived)

- This table is only relevant when the input option for processing is set to fixed or distribution.
- When the input option is set to fixed then in addition to the information in the first four columns only procnom or procupp needs to be specified. If both are specified the highest value will be used (worst case argument). For use in an uncertainty analysis also procnomuncupp may be specified.
- When the input option is set to distribution then in addition to the information in the first four columns procnom and procupp have to be specified describing the variability of processing factors. For use in an uncertainty analysis also procnomuncupp and/or procuppuncdf (or procuppuncupp[1]) may be specified.
- procupp should be higher than procnom.
- procnomuncupp should be higher than procnom.
- procuppuncupp should be higher than procupp.
- procuppuncdf should be positive, with values close to zero defining maximum uncertainty, and high values defining minimal uncertainty.
- Values lower than 0.01 are reset to 0.01; for processing types with disttype 1 (logistic) values higher than 0.99 are reset to 0.99.
- Procuppuncdf and procuppuncupp are alternative ways to specify uncertainty for the variability of processing factors[1]. The variability of processing factors is described by a standard deviation (at a logistic or logarithmic scale), and its uncertainty is described by setting the degrees of freedom (procuppuncdf) of a modified chi-squared distribution (see van der Voet and Slob, 2007 for an example). Alternatively, an upper uncertainty percentile on the upper variability percentile (procuppuncupp) can be specified, from which the appropriate number of degrees of freedom is derived by simulation[1].
- If a value for procuppuncdf is specified, procuppuncupp will be ignored.

1.7 Concentration tables

1.7.1 Compound (compulsory)

field name	Description
compound (text) compoundname (text)	compound code compound label (name of compound)

arfd (number)	ARfD (acute reference dose), in microgr/kg bw/day
adi (number)	ADI (acceptable daily intake), in microgr/kg bw/day
unit (number, optional)	-6 (default) or 0, -9, -12, -15 see below

- Missing values for ARfD and ADI: 9999.
- Column unit contains a coding to determine the unit as used for compound concentration data and dietary intake. Coding is as follows:

	concentration	intake
0	kg/kg	g/kg bw/day
-3	g/kg	mg.kg bw/day
-6	mg/kg	microgram/kg bw/day
-9	microgram/kg	nanogram/kg bw/day
-12	nanogram/kg	picogram/kg bw/day
-15	picogram/kg	femtogram/kg bw/day

- If column unit doesn't exist code -6 is assumed

1.7.2 Country (compulsory)

field name	Description
country (text)	code for country
countryname (text)	name of the country, label

- No missing values allowed

1.7.3 ConcentrationValues (compulsory)

field name	description
compound (text)	compound code
foodmeasured (text)	food code
year (number)	sampling year
month (number)	number of month
samplingtype (text)	type of sampling (monitoring)
country (text)	country of sample
numberofsamples (number)	count of the number of times the specified concentration or limit of reporting (LOR) occurs
value (number)	concentration (mg/kg) or LOR (see below)

- The limit of reporting is specified in column value using a minus (-) sign to make the distinction between a measured concentrations, *e.g.* –0.02 (see example first row).
- Concentration values are stored in column value and the number of times each value occurs in column numberofsamples, *e.g.* 0.21 and 1, respectively.
- Missing LORs are reported as –9999. The MCRA program replaces missing LORs with 1) the maximum LOR found in the database, 2) if all LORs are missing, the lowest concentration found in the database. A warning is generated when 1) and 2) are not possible.
- No missing values allowed for the other columns.

1.7.4 VariabilityProd (optional)

field name	description
food (text)	food code
varfac (number)	variability factor
coefvar (number)	coefficient of variation
nounitcomp (number)	number of units in the composite sample

- This table is used for specifying real empirical estimates of unit variability (*e.g.* from special studies) for the lognormal and the beta distribution and the number of units in a composite sample.
- Estimates for unit variability are independent of the compound.

- Missing values: 9999
- When the parameter for unit variability is a coefficient of variation and the number of units equals 1, unit variability is ignored for this food.

1.7.5 VariabilityCompProd (optional)

field name	description
compound (text)	compound code
food (text)	food code
varfac (number)	variability factor
coefvar (number)	coefficient of variation
nounitcomp (number)	number of units in the composite sample

- This table is used for specifying real empirical estimates of unit variability (*e.g.* from special studies) for the lognormal and the beta distribution that are dependent on the compound. Values for unit variability in table VariabilityProd are replaced by the new ones.

1.7.6 VariabilityProcCompProd (optional)

field name	description
compound (text)	compound code
food (text)	food code
proctype (number)	processing type code
varfac (number)	variability factor
coefvar (number)	coefficient of variation
nounitcomp (number)	number of units in the composite sample

- This table is used for specifying real empirical estimates of unit variability (*e.g.* from special studies) for the lognormal and the beta distribution that are dependent on the combination of processing type and compound. Values for unit variability in table VariabilityProd and VariabilityCompProd are replaced by the new ones. This can be used for example to reset the variability factor to 1 for grape juice and raisins (dried grapes).

1.7.7 AgriculturalUse (optional)

field name	description
compound (text)	compound code
food (text)	food code
country (text)	code for country
year (number)	year
useallowed (number)	indicator (0/1) whether use of the compound for the food is allowed (1) or not (0)
perccroptreated (number)	maximum percentage of the food that is treated with the compound

- For combinations of substance and foods that are not listed in table AgriculturalUse MCRA will assume that use is not allowed.

1.7.8 ConcentrationWorstcaseValues (optional)

field name	description
compound (text)	compound code
food (text)	food code
country (text)	code for country
year (number)	year
worstcasevalue (number)	worstcase value

- When information on detects and non-detects is missing, worstcase values may be used.