



Grant Agreement No: 245163

Project start date: July 2010

Project end date: July 2013

ACROPOLIS

Aggregate and Cumulative Risk Of Pesticides: an On-Line
Integrated Strategy

SEVENTH FRAMEWORK PROGRAMME

Deliverable 5.2: First prototype of the cumulative risk model.

Biometris unit of DLO Plant Research International,
part of Wageningen University and Research centre
Contact person: Hilko van der Voet, email: hilko.vandervoet@wur.nl

Project team: Hilko van der Voet, Waldo J. de Boer, Paul W. Goedhart, Gerie W.A.M. van der Heijden, Jaap Kokorian, Hilko van der Voet

Table of Contents

1	Introduction	4	
2	Context Diagram	5	
3	Use Case Specification	6	
4	Technical aspects	10	
4.1	Introduction		10
4.2	Model		10
4.3	View: User interface		11
4.4	Control: linking user interface and data model		13
5	Database Design and management	14	
6	References	15	

MCRA 8 - Implementation plan

20 June 2011

Background

DLO/Biometris will develop a new version of MCRA. There are several projects that require or ask us to implement new functionality:

- EU-Acropolis: multi-compound, cumulative exposure, aggregate exposure, more detailed uncertainty analysis, integration with effect modelling
- EFSA-ETUI: NCI method, 'First model, then sum' approach, use of FFQs e.g. to include never-consumer frequencies, importing an external within-person variance for usual intake
- EFSA-Guideline Probabilistic Modelling: loop over countries, parametric uncertainty approach for residue data

This implementation plan describes the intended adaptations to be made in MCRA. It tries to be as general as possible in order to accommodate all currently foreseen future developments in the structure of the program. Consequently, the content of this document should not be seen as describing guaranteed deliverables in a specific project.

1 Introduction

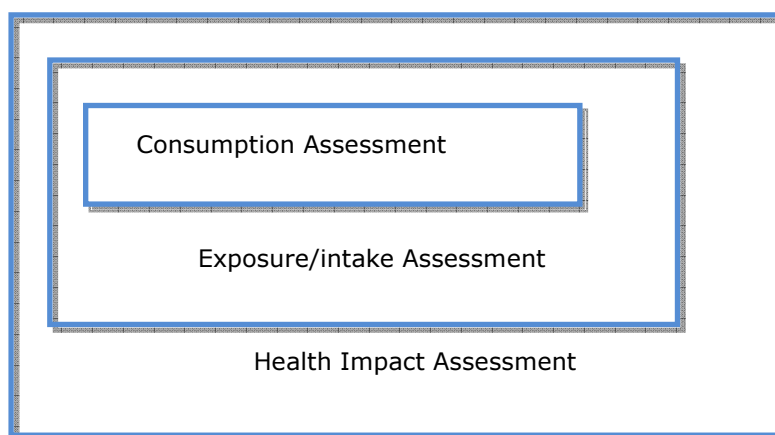
Chemical compounds in food can have health effects in a human population. Health effects can be negative ('risks') or positive ('benefits'). In addition to a dietary exposure there can also be a non-dietary exposure, e.g. dermal or by inhalation.

The MCRA software implements probabilistic methods for consumption assessment, exposure (or intake) assessment and health impact assessment for the use by risk (or risk/benefit) assessors, for example in the fields of pesticides, contaminants and additives, and assessors in the field of health and nutrition.

In the most wide case of use the assessor is interested in comparing scenarios for one or more chemicals occurring in one or more foods or other sources regarding their possibly cumulated and aggregated effect on one or more indicators of human health, either positive or negative. Probabilistic methodology for this general setting is termed probabilistic health impact assessment or PHIA, and in a specific implementation as integrated probabilistic risk assessment or IPRA (van der Voet et al. 2009).

Often the evaluation of health effects has been performed separately, and is summarised in human limit values such as the Acute Reference Dose (ARfD) or the Acceptable Daily Intake (ADI). In such cases only a probabilistic exposure assessment (equivalent term: intake assessment) is needed leading to a distribution of exposures (intakes) to be compared with the limit values.

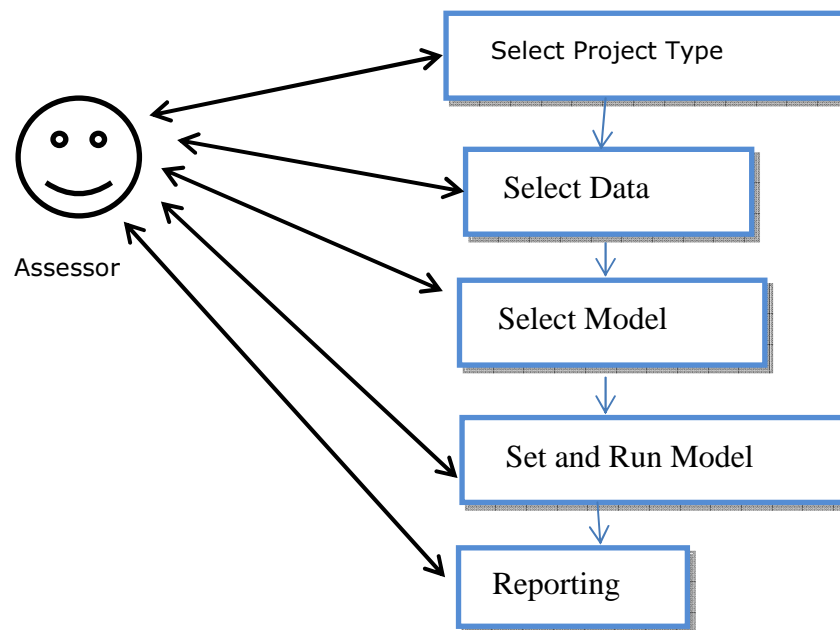
Sometimes also the concentration levels of the chemical (or nutritional component) in a single food are given as fixed values, and then only a probabilistic consumption assessment has to be performed.



Relevant background documents are: EFSA Guideline (2011, draft), van der Voet et al. (2009), de Boer et al. (2009), Boon et al. (2009).

2 Context Diagram

For this software one actor is defined “Assessor” this represents the scientist who uses the software to perform an assessment.



3 Use Case Specification

1. **Select Project Type:** the Assessor can select a Project Type combining the following possibilities (if appropriate):
 - a. The Assessment Type can be
 - i. Consumption Assessment: results are about amount of food consumed;
 - ii. Exposure Assessment: results are about exposure to (intake of) one or more compounds; or
 - iii. Health Impact Assessment: results are about margin of exposure relating to one or more compounds and one or more health effects.
 - b. For Exposure Assessments and Health Impact Assessments there are options
 - i. to indicate that there will be a Focal Food (as usual in e.g. pesticide registration assessments);
 - ii. to specify a Cumulative Assessment based on multiple compounds (which should be in a Common Assessment Group referring to a single health effect for Exposure Assessment);
 - iii. to specify an Aggregate Assessment that will combine non-dietary exposure and dietary exposure.
 - c. For Health Impact Assessments there is an option to specify the use of multiple health effects.
 - d. Health effect types can be Acute (short-term, in MCRA this is one day) or Chronic (long-term).
 - e. The basic data type can be
 - i. Consumption data at PersonDay level plus (for Exposure and Health Impact Assessments) Concentration data per food
 - ii. Exposure data at PersonDay level
 - iii. For Health Impact Assessments there should be additional data on doses that have critical health effects.

There is an option to generate the second type of data by simulation.

- f. The number of assessments to be executed in a single run can be
 - i. Single assessment;

- ii. Two assessments. There can be options to specify predefined choices, e.g. specify the EFSA PPR Guideline optimistic and pessimistic scenarios;
- iii. Multiple assessments. There can be options to specify predefined choices, e.g. assessments by Country.

In the case of more than one assessments, an assessment group is created, and pivotal results will be tabulated across the assessments.

Several Default Project Type choices will be available, e.g. for acute con-cumulative, non-aggregate exposure assessment, etc. Starting from an existing Project it will be possible to start a wizard to create a new Project with adapted Project Type.

Project Type selection

Assessment Type

<input type="radio"/> Consumption Assessment				
<input type="radio"/> Exposure Assessment	<input type="checkbox"/> Focal Food	<input type="checkbox"/> Cumulative	<input type="checkbox"/> Aggregate	
<input type="radio"/> Health Impact Assessment	<input type="checkbox"/> Focal Food	<input type="checkbox"/> Cumulative	<input type="checkbox"/> Aggregate	<input type="checkbox"/> Multiple health effects

Health Effect Type

<input type="radio"/> Acute (short-term)
<input type="radio"/> Chronic (long-term)

Data Type

<input type="radio"/> Read Consumption + Concentration
<input type="radio"/> Read Exposure at PersonDay
<input type="radio"/> Simulate Exposure at PersonDay

Number of Assessments

<input type="radio"/> Single Assessment			
<input type="radio"/> Two Assessments	<input type="checkbox"/> EFSA Guidance	<input type="checkbox"/>	<input type="checkbox"/>
<input type="radio"/> Multiple Assessments	<input type="checkbox"/> by Country	<input type="checkbox"/> Repeat simulations	<input type="checkbox"/>

- Set current choices as default
- Skip this screen next time and use default

2. **Select data:** the Assessor can use data from one or more data areas including 'central' (pointing to a directory on the MCRA website server) and 'own' (referring to a directory on the Assessor's own computer).

Data can be in Access or Excel. An alternative is to generate simulated data. Simulated data can be stored as one or multiple Excel files.

With Access data he can select tables from one or multiple databases. With Excel data he can select one or more Excel files.

More than one table of the same type can be selected by selecting a directory containing multiple versions of the same type. Depending on the selected Project Type this will be interpreted to indicate that the assessment has to be done e.g. “by scenario” or “by country”, if appropriate. Otherwise multiple data tables of the same type will be appended.

Alternatively, in Project groups with two or more assessments the data are first selected for the first Project. Based on this project, it will be possible to change data selections (and model selections) for the other projects in the group.

All data will be validated for appropriateness in the chosen scenario (e.g. multi-compound needs sample-based concentrations).

A translation from food-as-eaten to food-as-measured is made.

The user can perform further subset selections, regarding compound(s), individuals, foods (as-eaten and as-measured).

The user can inspect or edit the selected data (for centrally provided data only if data access options according to database manager set options allow so). Edited data can be saved in the ‘Own’ data area.

3. **Model concentration data:** the Assessor can select, food by food, a model for the concentration data (including non-detects). This includes options for processing factors, unit variability, and pesticide usage data.

For the multi-compound acute scenario the user can select a model for cumulative exposure (e.g. Stochastic EM or Simple).

4. **Select model:** the Assessor can select detailed modelling options for the acute or chronic exposure models. In both scenarios the use of covariates (e.g. age, gender of FFQ answers) can be specified.

The user can specify the number of iterations for Monte Carlo calculations, and provide a seed for the random generator to ensure reproducible calculations.

For chronic risk assessment options include the type of model for chronic risk (Observed Individual Means, BetaBinomial-Normal, LogisticNormal-Normal, DiscreteSemiParametric), the level of modelling (single food or aggregated foods), the inclusion of never-consumers (external fraction), or the inclusion of an external variance for within-person variation.

5. **Uncertainty analysis:** the Assessor can specify whether or not to perform a quantitative uncertainty analysis, and choose the number of iterations. The user can turn on or off the inclusion of specific sources of uncertainty (such as Individuals and Portion size for consumption data, Samples and Processing factors for

concentration data), and can ask for an analysis to quantify the contributions of the uncertainty sources to the total uncertainty (sensitivity analysis).

The user can ask to fill a table for non-quantified (qualitative) uncertainties.

6. **Set and Run:** the Assessor can specify details of the output to be calculated. This includes:
 - a. which percentiles or exceedance limits to calculate (either as fraction of ArfD or ADI, or as explicit values)
 - b. levels of covariables where the exposure will be calculated.
 - c. A percentile to specify drill-down possibilities

The user can start the main computation which will be processed off-line.

7. **Reporting:** the Assessor can view or save reports generated from the statistical analysis
 - a. The user is able to view the results of the statistical analysis and has access to the standard printing options available in Windows applications.
 - b. The user can zoom in on the right tail of plots, and adapt the scale (e.g. logarithmic or not)
 - c. The user can perform a drill-down on selected individuals showing the detailed calculations underlying the final value.
 - d. The user is able to save the results of the statistical analysis in either .pdf or .rtf format.

4 Technical aspects

4.1 Introduction

This chapter describes several technical aspects of the implementation of MCRA 8. The following topics are treated in order: the logical class structure, the graphical user interface and the link between the two. This structure roughly corresponds to a Model-View-Control, or Model-View-ViewModel software architecture.

4.2 Model

4.2.1 Introduction

A complete MCRA risk assessment, regardless the exact details of the kind of analysis, can be described by a collection of items that fall in three categories:

- input data
- settings
- output data

Combined, these items form a 'Project'. A project has an unique name and is owned by a specific 'User' (the Assessor). Each project may have multiple 'settings', that are associated with different logical parts of a simulation.

4.2.2 Projects Class implementation

A Project is represented by an instance of the Project class. The Project objects are read from and written to by the executable that runs the actual calculation (MCRA.exe), they are filled with settings through the user interface, and read by output visualizers. In other words, they are the primary container for all data that is associated with a specific simulation. Projects should be serializable so they can be exchanged with other tiers over TCP/IP connections and should be saved in the backend database, so they can be recalled in a different user session.

4.2.2.1 Input data

After a user selects input source databases, the data is read, and attached to the project by means of Entity Data Model objects. This ensures that during the user session in which the project is created, the data is fully loaded in memory and can be queried directly. When the user session ends, the input data is discarded, but a reference to the source databases is kept. This way, the data can be reloaded in a future user session, provided the data sources are still available.

4.2.2.2 Intermediate data

At some point during a project, the input data is converted to a different form. In MCRA 7, this conversion is performed by an external executable (Conversion.exe), this may be changed to a direct conversion from the web interface. The results of the conversion will be added to the project's input data objects (not to the original data-sources). When a user-session ends, the converted data of the most recently executed project is saved in the backend database.

4.2.2.3 Settings

The settings for a MCRA simulation are subdivided in logical settings groups, each corresponding to a certain aspect of the simulation. For example, there is a `ModelSettings` class, that could have properties such as:

- `ModelType`: enum { OIM, BBN, LNN, DSP }
- `Level`: enum { Single, Aggregated }
- `MonteCarloIterations`: integer
- `RandomSeed`: integer
- `IncludeNeverConsumers`: Boolean
- `IncludeExternalVariance`: Boolean

These settings correspond to options that have to be set by the user. There is also a settings class that is associated with uncertainty analysis, input data conversion, concentration modeling.

Settings can be saved independently of a project, allowing for templates to be created and loaded for each logical settings group.

4.2.2.4 Output data

The results of a MCRA simulation are stored in the project's output field(s) in raw form, and are saved in the backend database, so they are instantly available in future user sessions.

4.2.3 Project groups for batch processing

Projects can be grouped in 'project groups'. Project groups provide two types of functionality:

- Batch execution
- Combining project outputs

Key to the concept of project groups is the fact that the projects within the group are 'unaware' of the fact that they belong to a group. The project group has access to all its members, but they do not have access to each other and behave exactly as usual.

4.2.3.1 Batch execution

The user sets up several projects with different settings and or input data, groups them together and executes them all at once.

4.2.3.2 Combining project outputs

The project group exposes the output of its member projects in aggregated form, and adds some additional data, such as inter-project averages and variances. This allows an output-visualizer to show multiple graphs in one image.

4.3 View: User interface

In this section an overview of the important features and components of the user interface is given.

4.3.1 Interface philosophy

The objective of the interface design is that users can create new projects without having to think about "which settings he still has to set", because he is guided through all the steps by a wizard. With the same structure, the more confident user can setup a new project, or make changes in an existing project, by means of a non-linear, flat view that he can access at will.

The user interface consist of two main component categories:

- **Overviews** present a certain underlying data source object to the user in a one-way fashion: the user can only look at it.

- **Editors** are the interactive version of an overview and allow the user to edit the properties of the underlying data source object.

Every overview and every editor corresponds to a single Settings class in the underlying data-model.

The user interface implements two different ways of setting up a project:

- **Linearly**, by means of a wizard that guides through all the steps required from start to end. The wizard consists of a number of 'pages', or 'steps' that confront the user with an Editor object and some explanation that corresponds to the Settings object that is being edited.
- **Non-Linearly**, by means of a tabbed view. The tabs carry the (full) names of the settings classes and contain an Overview component of the corresponding settings class. By pressing an "Edit" button next to the overview, a dialog pops up that shows the Editor object for this Settings class.

Key to the implementation of this concept is that the Overview and Editor components are designed in such a way that they can be used in both a wizard and a tab-panel. This clearly separates the design of the main user interface, the design of the wizard state-machinery and the Overview and Edit components themselves, greatly simplifying project planning and management.

4.3.2 Main view

A prototype for the main screen of the user interface is shown in Figure 1. It shows all projects owned by the user in a TreeView and displays the contents of the selected project in a tabbed panels. Projects can be organized in folders (without additional functionality) and in project groups, that provide the functionality described in section 4.2.3.

Next to the project tree is the previously mentioned tabbed overview of the project's settings. Besides tabs that correspond to settings objects, there is also an "overview" tab that shows a summary of the most important settings, as well as the project's creation date and execution date.

The last tab contains an overview of the output results as well as buttons that offer the user several download options (raw data, pdf-report, etc.).

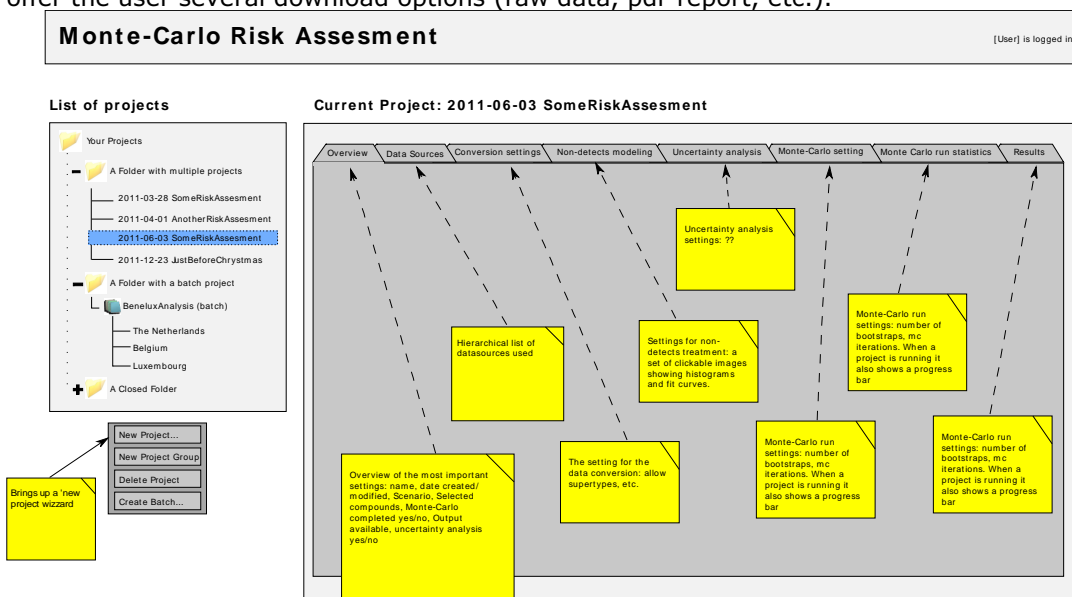


Figure 1: Main view of the MCRA client application.

4.3.3 Create project wizard

When the user creates a new project, the "New Project Wizard" pops up, guiding the user through the steps that are outlined in section 3. Depending on the

selections the user chooses during the first step (see prototype in Figure 2), the available settings that have to be set during the next steps of the wizard become limited. Also, different tables of input data are required, some become optional and some become mandatory. A prototype for the screen that lets the user select the input data is shown in Figure 3.

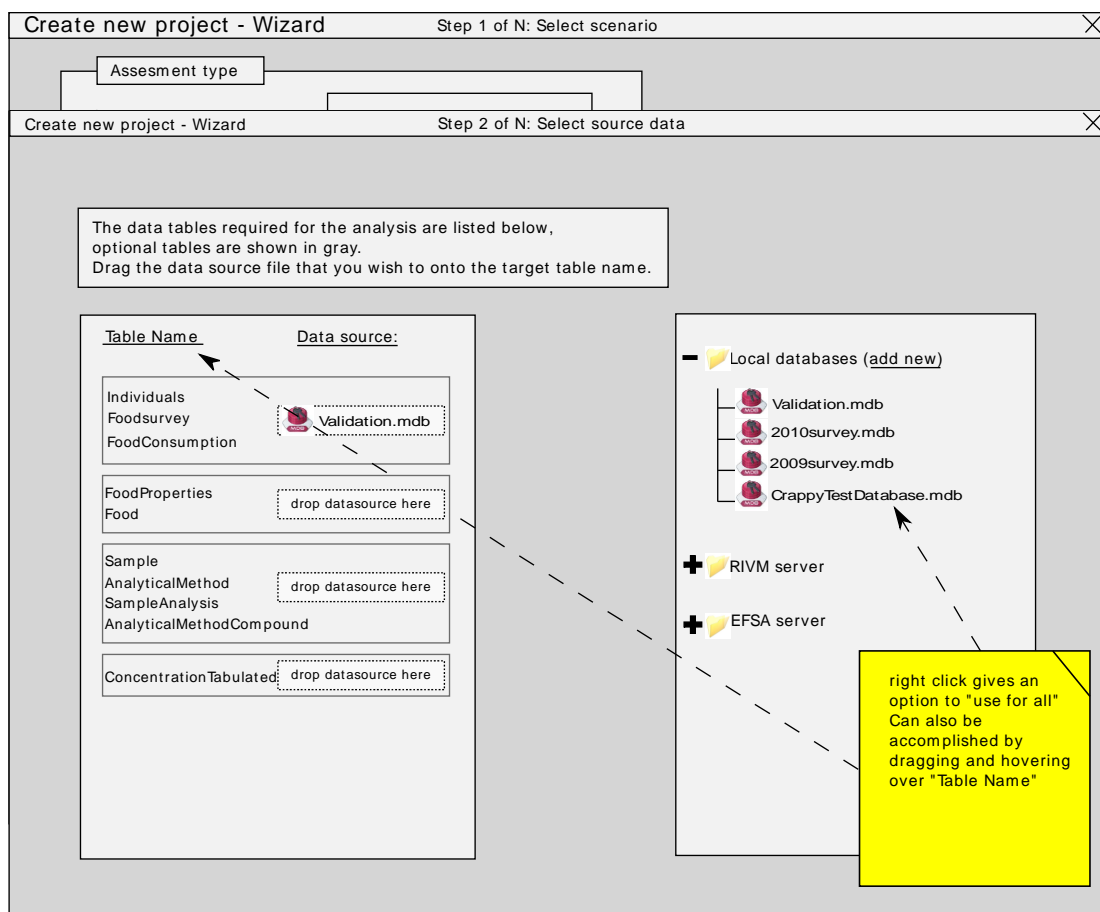


Figure 3: Step two of the "New Project wizard". The user can select which data sources he uses for each group of input tables.

4.3.4 Batch project wizard

The batch project wizard helps the user to duplicate his project multiple times and group all the duplicates in a Project Group. The wizard offers a choice of settings (in theory any setting, including input data selection) that the user can choose to vary. For example, the user can choose the setting "CoVariate" to be set to all possible values ("gender" and "age"). The projects in the project group that is created by the wizard are now fully equal, except for this one setting, that is set to "gender" in one project and to "age" in another.

The wizard also provides functionality to allow the user to run a project with the same settings, but with different data sources, by letting him select a folder that contains multiple source data files.

4.4 Control: linking user interface and data model

The link layer between the User Interface (view) and the underlying Projects-Settings class structure (model) exposes the objects in the model layer to the view layer, so the View's component can interact with them in such a way that

no business logic is required in the view layer itself. Its most important responsibility is data validation, but it also contains classes that, for example:

- convert a projects raw output data objects to tables and figures, complete with headings, labels and tooltips
- ...

5 Database Design and management

5.1 Introduction

There are several databases that have to be maintained:

- The main backend database. This contains the saved projects and user information. It also contains calculation results (related to projects). The database is located on the MCRA database server.
- 'User owned database files.' This is the collection of Access database files and Excel spreadsheet files that a user has uploaded to the MCRA server to his user area from his local PC for use as a data source.
- (optional) 'Company owned database files'. This is a similar collection of Access database files and xls files as in the 'user owned database', that is located on a remote server.
- (optional) 'Company owned database servers'. This is a relational database on a remote server that offers data to other users organized according to the format specified for user-owned Access database files.

Depending on the way the user interface is implemented, the management of the databases is either the full responsibility of the MCRA website server, or the shared responsibility of the client application.

5.2 Data exchange between tiers

5.2.1 User data

The data from these tables is 'shaped' into a different form by the conversion algorithm. If possible, the conversion algorithm will be performed client-side, at the same time the source data files are read. The data will then be available in the form of Entity model class instances, that are shared with the MCRA website server with the aid of windows communication foundation. The backend database is only used for data persistence between user sessions. During a session, all data is available in the form of Entity Data Model objects and can be queried using LINQ statements.

5.2.2 Company database servers

Any data on a remote company server can be read by the website server using ADO.NET or entity framework classes. The data can be used by the user, but is never actually present on the clients computer. A clever way has to be found to deal with the scenario where the user wishes to do an analysis of his own data combined with that on a company server. When dealt with naively, it can result in huge amounts of data being transferred back and forth between the client and the MCRA website server.

5.3 Database design

Datasources provided by the user contain the following tables:

- Compound
- FoodSurvey
- Individual
- Food
- FoodProperties

- AnalyticalMethod
- Sample
- SampleAnalysis
- ProcessingType
- AnalyticalMethodCompound
- RelativePotencyFactor
- ConcentrationTabulated
- ConcentrationPerSample
- ProcessingFactor
- FoodConsumption
- FoodTranslation
- Effect
- DoseResponse
- CriticalEffectDose
- CEDUncertaintySample
- AgriculturalUse
- VariabilityFactor
- FoodMarketShare
- FoodConsumptionQuantification
- FoodConsumptionUncertainty

We will refer to the data in these tables as the 'source data'. The table relationships can be shown in a diagram like Figure 4. The user provides these tables in Microsoft Excel (xls) or Microsoft Access (mdb) format.

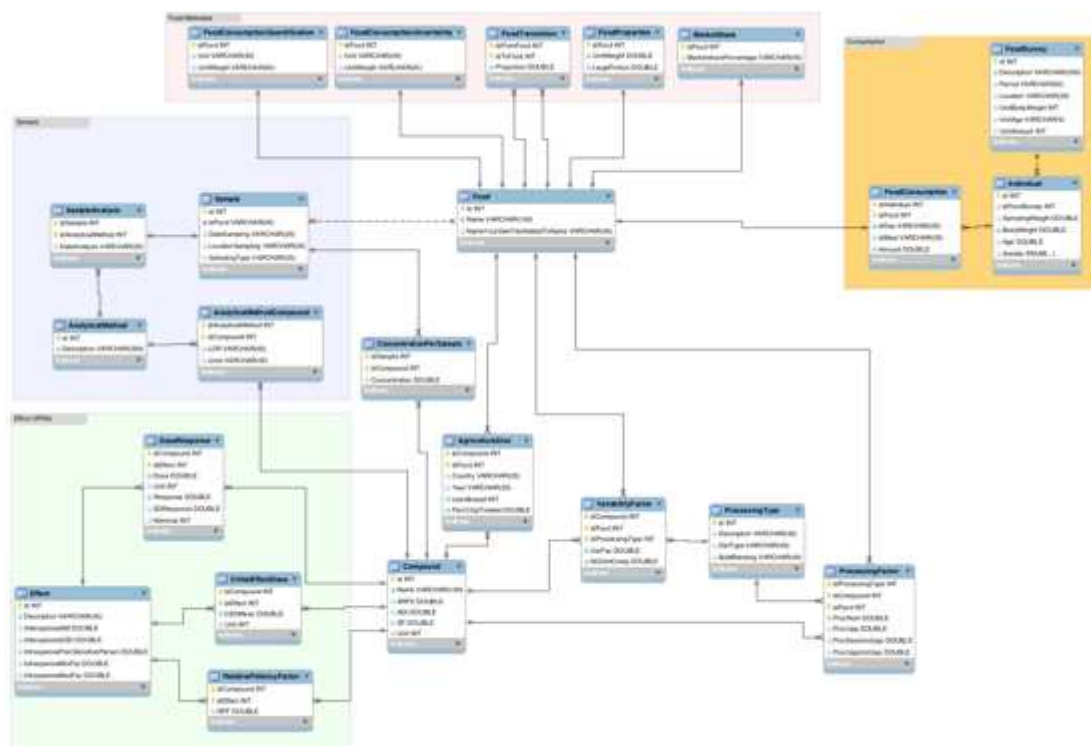


Figure 4 The input source database tables and their relations

6 References

Boon, P.E., Svensson, K., Moussavian, S., van der Voet, H., Petersen, A., Ruprich, J., Debegnach, F., de Boer, W. J., van Donkersgoed, G., Brera, C., van Klaveren, J. D. and Busk, L. (2009). Probabilistic acute dietary exposure assessments to captan and tolyfluanid

using several European food consumption and pesticide concentration databases. *Food and Chemical Toxicology*, 47: 2890-2898.

de Boer, W.J., van der Voet, H., Bokkers, B.G.H., Bakker, M.I. and Boon, P.E. (2009). A comparison of two models for estimating usual intake addressing zero consumptions and non-normality. *Food Additives and Contaminants. Part A – Chemistry Analysis Control Exposure & Risk Assessment*, 26: 1433-1449.

van der Voet, H., van der Heijden, G.W.A.M., Bos, P.M.J., Bosgra, S., Boon, P.E., Muri, S.D. and Brüscheweiler, B.J. (2009). A model for probabilistic health impact assessment of exposure to food chemicals. *Food and Chemical Toxicology*, 47: 2926-2940.