Downloaded from orbit.dtu.dk on: Dec 20, 2017

Technical University of Denmark

Development of a user-friendly interface version of the Salmonella source-attribution model

Hald, Tine; Lund, Jan

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Hald, T., & Lund, J. (2012). Development of a user-friendly interface version of the Salmonella source-attribution model. Søborg: The National Food Institute, Technical University of Denmark.

DTU Library
Technical Information Center of Denmark

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

EXTERNAL SCIENTIFIC REPORT

Development of a user-friendly interface version of the *Salmonella* source-attribution model[1]

Tine Hald[a] and Jan Lund[b][2]

[a] National Food Institute, Technical University of Denmark, Mørkhøj, Denmark

[b] NextPhase IT, Jersie, Denmark

ABSTRACT

The objective of the work described in this report was to develop a flexible and user-friendly interface for attributing human cases of food-borne pathogens to the responsible food-animal reservoirs and/or food sources. The interface is based on two existing *Salmonella* source-attribution models developed in previous EFSA contracts. The developed interface is called the EFSA Source Attribution Model (EFSA_SAM) and the programming language used is Embarcaderos Delphi XE2 Enterprise. Based on the user's imported data and model selections, the interface generates a WinBUGS code that is executed in WinBUGS and the resulting data are then imported from WinBUGS to the interface software for tabulation and graphical display. This approach ensures consistency in both model and data setup, eliminating the need for user knowledge of WinBUGS syntax. EFSA_SAM requires data by country on reported number of human cases by subtypes, food-source prevalences by subtypes and food production and trade. Users can specify which countries, food sources and subtypes (e.g. *Salmonella* serovars) to include in the model. The EFSA_SAM also includes the possibility to run different scenario analyses, where the user can explore the effect on human cases by changing the prevalence of specific subtypes in the included food sources. A critical part of all Bayesian models is to check for model convergence and goodness of fit. The EFSA_SAM describes different ways for checking convergence and exploring model fit providing the user with some tools to assess the validity of the model. The EFSA_SAM interface is delivered with a user-manual, which is also part of this report. Users of the interface are recommended to read this report before starting using the interface to become familiar with the model principles and the mathematics behind.

© National Food Institute, Technical University of Denmark

KEY WORDS

Salmonella, source attribution, flexible user-friendly interface version, scenario analysis

DISCLAIMER

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

[1] Question No EFSA-Q-2012-00672.

[2] Acknowledgment: The authors would like to thank EFSA staff Winy Messens, Luis Vivas Alegre (BIOHAZ Unit) and José Cortinas Abrahantes (SAS Unit) for the support provided to this external scientific report submitted to EFSA.

Any enquiries related to this output should be addressed to biohaz@efsa.europa.eu. The EFSA_SAM model is available upon request to biohaz@efsa.europa.eu.

Suggested citation: Tine Hald and Jan Lund; Development of a user-friendly interface version of the *Salmonella* source-attribution model. Supporting Publications 2012:EN-318. [77 pp.]. Available online: www.efsa.europa.eu/publications

© European Food Safety Authority, 2012

SUMMARY

EFSA has been working on a series of Scientific Opinions originating from a mandate received by the European Commission (EC) in July 2008 on the review of *Salmonella* targets in poultry primary production. For evaluating targets in the broiler and turkey production, specific *Salmonella* source attribution models have been developed by external contractors. Both models were based on the Hald model and use a Bayesian approach employing microbial subtyping data, in both cases *Salmonella* serovar data. These types of source attribution models allow for the identification of the most important animal reservoirs of the zoonotic agent, assisting risk managers to prioritize interventions and focus control strategies at the animal production level. The model can provide estimates for the effect on the number of human cases originating from a particular reservoir, if the observed prevalence in that reservoir is changed or for specific subtypes e.g. specific serovars of *Salmonella* occurring in that reservoir.

The source-attribution approach has been considered by EFSA Working Groups and Panel Experts as valid when addressing these types of questions, where the use of a classical quantitative risk assessment model (i.e. transmission models) would be impaired due to a lack of data and time limitations. As these models require specialist knowledge, it was requested by EFSA to develop a flexible user-friendly source attribution model for use for example in future mandates dealing with similar questions.

The objective of the work described in this report was, therefore, to develop a flexible and user-friendly interface for attributing human cases of food-borne pathogens to the responsible food-animal reservoirs and/or food sources. The interface is based on a *Salmonella* source-attribution model developed for setting target for *Salmonella* in the turkey production: the Turkey Target Source Attribution Model (TT-SAM). Results from this model were used by the BIOHAZ panel in their related Scientific Opinion. The developed interface described in this report is called the EFSA Source Attribution Model (EFSA_SAM).

The programming language (development environment) used for developing the user-friendly interface is Embarcaderos Delphi XE2 Enterprise. The interface generates a WinBUGS code based on the user’s imported data and model selections. The interface exports this code with corresponding data to WinBUGS, where the code is executed automatically. The model results are then imported from WinBUGS to the interface software for tabulation and graphical display, and possible exportation to other softwares for further analysis e.g. MS Excel. This approach ensures consistency in both model and data setup, eliminating the need for user knowledge of the WinBUGS syntax.

Users can import data into the EFSA_SAM from semicolon-separated files. Required data are i) the reported number of human cases per country and subtype including data on the number of travel and outbreak-related cases, also per country and subtype, ii) food-animal prevalence data per country and subtype, including the number of units tested and the number of positive units, and iii) data on the production and trade of the different food-animal sources in the EU Member States. The EFSA_SAM also allows for the inclusion of underreporting factors recognizing that the reported number of human cases only reflects a part of the disease burden and the degree of underreporting varies hugely between countries. In the interface users can specify which countries, food sources and subtypes (e.g. *Salmonella* serovars) to include in the model.

It is also possible to run an analysis for a single country only, but where several periods (typically years) of data are included. This can provide an indication of the trend over time. Required data for this type of model are i) the reported number of human cases by subtype including data on the number of travel, domestic, unknown travel history and outbreak-related cases, also per subtype, ii) food-animal prevalence data per subtype, including the number of units tested and the number of positive

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

units, and iii) data on the amount of the included animal foods available for consumption in the country.

The data imported into EFSA_SAM will be used for a baseline analysis providing estimates on the number of human cases attributable to the different food-animal sources in the actual situation. The results of the baseline analysis can be compared with the results from one or more scenario analyses specified by the user. The interface allows for two types of scenarios: i) the setting of target prevalences for individual subtypes, and ii) the setting of a combined target prevalence for a group of subtypes. In the first type, EFSA_SAM will automatically change the original prevalence to the set target prevalence, but only if the original prevalence is greater than the target prevalence. In the latter type, the users can select any number of subtypes for which a combined prevalence should be equal or less to a set target prevalence. The EFSA_SAM generates a new set of subtype-specific prevalences that are proportionally scaled down from the original prevalences in order to result in an overall prevalence equal to or less than the target prevalence. A comparison of the baseline and scenario results can be used to assess the effect on the predicted number of human cases, if targeted control measures are implemented for specific subtypes or groups of subtypes.

A critical part of all Bayesian models is to check for model convergence and goodness of fit. The EFSA_SAM describes different ways for checking convergence and include the calculation of the bgr-diagnostics that is also a part of the WinBUGS software. For exploring goodness of fit, a ratio between the observed and predicted number of human cases per country is calculated. A poor fit of the model for some countries is often linked to poor data quality.

The EFSA_SAM interface is delivered with a user-manual, which is also part of this report. Users of the interface are recommended to read this report before starting using the interface to become familiar with the model principles and the mathematics behind, which is required in order to interpret the model results and assess the validity of the model.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

TABLE OF CONTENTS

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

BACKGROUND AS PROVIDED BY EFSA

EFSA has been working on a series of Scientific Opinions originated by a mandate received by the European Commission (EC) in July 2008 on the review of *Salmonella* targets in poultry primary production. The Opinions have been adopted by the BIOHAZ Panel and published on the EFSA website. They have provided a quantitative estimate of the public health impact of setting new targets for the reduction of *Salmonella* in breeding flocks, laying hens, and broilers of the species *Gallus gallus*.

A similar mandate for flocks of breeding and fattening turkeys was received by EFSA in June 2010 (EFSA-Q-2010-00899) and this Opinion was published in April 2012. EFSA was asked by the EC to indicate and rank the *Salmonella* serotypes with public health significance according to Annex III of Regulation (EC) No 2160/2003[3], to assess the impact of a reduction of the prevalence of *Salmonella* in breeding flocks of turkeys on the prevalence of *Salmonella* in flocks of fattening turkeys and to assess the relative public health impact if a new target for reduction of *Salmonella* is set in fattening turkeys being 1 % or less of flocks remaining positive for all *Salmonella* serovars with public health significance compared to:

- the theoretical prevalence at the end of the transitional period (1 % or less of flocks remaining positive for *Salmonella* Enteritidis or *Salmonella* Typhimurium), and
- the real prevalence in 2010 to be reported by the Member States (MSs).

The three Opinions addressing *Gallus gallus* have employed a different approach to address the quantitative aspects of the questions received. The Opinion related to *Salmonella* in broilers was supported by the work of a contractor (CT/EFSA/BIOHAZ/2010/02) who provided quantitative estimates using a broiler-target *Salmonella* source-attribution model (BT-SAM). This model was based on the Hald model and uses a Bayesian approach employing microbial subtyping data (Hald et al., 2004). The Hald model is a Markov Chain Monte Carlo model written in the WinBUGS environment on a PC. The Contractor’s model uses the latest version, called OpenBUGS (see http://www.openbugs.info/w/), and has added the EU MSs to the model as a third dimension. This type of model allows for the identification of the most important reservoirs of the zoonotic agent, assisting risk managers to prioritize interventions and focus control strategies at the animal production level. The model can provide estimates for the effect on the number of human cases originating from a particular reservoir, if the observed prevalence in that reservoir is changed or for specific subtypes e.g. specific serovars of *Salmonella* in that reservoir.

This source-attribution approach has been considered by WG and Panel Experts as valid when addressing this type of questions, where the use of a classical quantitative risk assessment model (i.e. transmission model) would be impaired due to a lack of data and time limitations. Therefore, for the Opinion related to *Salmonella* in turkeys, a similar model, named the turkey-target *Salmonella* source-attribution model (TT-SAM) was used to support the BIOHAZ Panel (see previous contractor’s report: Hald et al., 2012).

The BT-SAM model, developed by the former Contractor is written in an OpenBUGS code, and the TT-SAM model is written in WinBUGS 1.4.3. As these models require specialist knowledge, it was requested by EFSA to develop a flexible user-friendly source attribution model for use for example in other mandates dealing with similar questions.

[3] Regulation (EC) No 2160/2003 of the European Parliament and of the Council of 17 November 2003 on the control of *salmonella* and other food-borne zoonotic agents. OJ L 325, 12.12.2003, p. 1-15.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

TERMS OF REFERENCE AS PROVIDED BY EFSA

The purpose of the contract was to provide the EFSA BIOHAZ panel with a source attribution model to evaluate targets specifically answering the Terms of Reference of the *Salmonella* in turkey flocks mandate (EFSA-Q-2010-00899), which has been covered in a previous report (Hald et al., 2012). In addition, the objective is to provide EFSA with a "generic" flexible, but at the same time user-friendly interface for source-attribution modelling of food-borne pathogens, based on the codes of the existing *Salmonella* source-attribution model. This interface can be explored to evaluate the public health impact of different prevalence levels and target serovars of *Salmonella* spp. in fattening turkey flocks in EU MSs. Appropriate training to future users of the model should also be provided.

According to the Technical Specifications of the Negotiated Procurement Procedure NP/EFSA/BIOHAZ/2011/04, the tasks to be carried out in particular are:

to develop a flexible user-friendly "generic" interface for a source-attribution model.

- The broiler-target *Salmonella* source-attribution model (BT-SAM) should be employed to develop a windows platform user-friendly "generic" and flexible model interface.

- The model (which is now 3-dimensional) should allow for flexibility. It should be possible to change the input parameters of the model, i.e. the number of animal-food sources, the number of MSs, and the type and number of subtypes.

- It should be possible to evaluate various scenarios using the user-friendly interface. It should allow to change the prevalence of single subtype in one or more animal-food sources and the combination of different subtypes in one or more sources (e.g. for evaluating targets).

- The user-friendly interface should allow for generating outputs for different scenarios: the expected rates of human cases in the EU MSs (including underreporting) and the percentages in terms of the EU expected rate (mean statistics, 2.5 % and 97.5 % statistics) by food type, subtype as done in the previous contractor's report. The interface should include some graphical tools as well (e.g. density plots of the uncertainty distributions of the total reduction of human cases of foodborne-related salmonellosis and these related to a targeted animal-food source only for the scenarios). A menu should be provided to include only a selection of the data, for example for incorporating only one MS and excluding/including trade. The interface should also have a reporting menu, in which results to be used are output and presented in a sort of report format.

This contract/grant was awarded by EFSA to:

The National Food Institute, Technical University of Denmark, Denmark

Contract title: Development of a flexible user-friendly interface version of the *Salmonella* source-attribution model developed under CFT/EFSA/BIOHAZ/2010/02 for evaluating targets in turkey meat production (EFSA-Q-2010-00899) and use in future source-attribution assessments.

Contract number: CT/EFSA/BIOHAZ/2011/02

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

OBJECTIVES

The overall objectives of the tasks covered by this report were:

- To develop a flexible and user-friendly interface for attributing human cases of *Salmonella* to responsible food-animal reservoirs and/or food sources. The interface is based on a *Salmonella* source-attribution model developed for setting target for *Salmonella* in the turkey production (the Turkey Target Source Attribution Model: TT-SAM), which has been described in a previous report (Hald et al., 2012). The TT-SAM model was based on two existing *Salmonella* source-attribution models developed in the WinBUGS software as part of previous EFSA service contracts (CFT/EFSA/BIOHAZ/2010/02 and NP/EFSA/ZOONOSES/2010/01).

- To prepare a user-manual explaining the use of the interface software.

MATERIALS AND METHODS

1. Principle of the Bayesian subtyping approach for source attribution modelling

The microbial subtyping approach involves characterisation of isolates of the pathogen by phenotypic and/or genotypic subtyping methods. The principle is to compare the distribution of subtypes in potential sources (typically food animals) with the subtype distribution in humans, and the approach is enabled by the identification of strong associations between some of the dominant subtypes and a specific food-animal reservoir, providing a heterogeneous distribution of subtypes among the sources. Subtypes exclusively or almost exclusively isolated from one source are regarded as indicators for the human health impact of that particular source, assuming that all human infections with these subtypes originate only from that source. Human infections caused by subtypes found in several reservoirs are then distributed relative to the prevalence of the indicator types.

The Bayesian model first described by Hald et al. (2004) attributes domestically acquired laboratory-confirmed human infections caused by different *Salmonella* subtypes (e.g. serovars, phage types, antimicrobial resistance profiles) as a function of the prevalence of these subtypes in animal and food sources and the amount of each food source consumed. However, the number of people being infected by a particular subtype in a particular food source supposedly depends on additional factors related to the subtype and food source in question. Therefore, a multi-parameter prior, which accounts for the presumed but undefined differences between subtypes and food sources with respect to cause human infections, was introduced.

The bacteria-dependent factor $\{q_i\}$ can be interpreted as combining survivability, virulence, and pathogenicity of the pathogen to estimate the ability of that subtype to cause disease, whereas the food source dependent factor $\{a_j\}$ estimates differences between food sources in characteristics that affect their ability to act as vehicles for foodborne infections (e.g. general differences in bacterial load, food characteristics influencing growth behaviour, or preparation procedures). It is, however, emphasised that the estimated values of the bacteria- and food-source-dependent factors are simply multiplication factors (comparable to regression coefficients in regression analyses) that helps us to arrive at the most likely solution given the observed data. Their relative size can provide an idea about the differences between subtypes and food types with respect to causing human infections, but estimates based on the results of a single model should be interpreted with care. However, by applying the model on a regular basis as new data become available, it may be possible to monitor the main sources and dynamics in the occurrence of human salmonellosis and to improve the estimation of the model parameters, including the bacteria- and food-source-dependent factors.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

The basic equation used to estimate the number of human cases per source and subtype is defined as follows:

$$\lambda_{ij} = p_{ij} * M_j * a_j * q_i \qquad \text{Equation 1}$$

where λ_{ij} is the expected number of cases per subtype *i* and source *j*, p_{ij} is the prevalence of subtype *i* in source *j*, M_j is the amount of source available for consumption in the country, a_j is the food-source dependent factor for source *j*, and q_i is the bacteria dependent factor for subtype *i*. To avoid problems related to identifiability (i.e. overparameterisation) of the model described in Eq. 1, the number of estimated parameters needs to be reduced. The pooling of some subtypes or food sources into groups with similar characteristics is one way of addressing this problem. Depending on the available data, the model can be extended to include other dimensions such as time period (e.g. year) and country, which can also increase the robustness of the model and consequently improve the parameter estimation for instance by assuming that the q-values remain unchanged over at least shorter time periods (Pires and Hald, 2010) and are independent on country.

The model calculates the expected number of human cases per subtype $\{\lambda_i\}$ according to the above equation. From this λ_i, a back-calculation is made by adding the number of travel- and outbreak-related cases with known subtype in order to get the expected number of reported cases. The observed data (i.e. the reported number of human cases per subtype) is then linked with the prior distribution by assuming that the number of cases per subtype is Poisson distributed (the likelihood function) with a parameter value equal to the expected number of cases. This results in posterior estimates for the unknown parameters q_i and a_j and consequently for the number of cases per subtype and source $\{\lambda_{ij}\}$, which can then be summarised over subtypes to get to the number of cases per source $\{\lambda_j\}$.

The microbial subtyping approach requires a collection of temporally and spatially related isolates from various sources and humans, and is consequently facilitated by an integrated food-borne disease surveillance programme focused on the collection of pathogen isolates from the major food animal reservoirs and from humans (Pires et al., 2009). The data quality and availability are considered the biggest limitation of this approach.

A strong advantage of the microbial subtyping approach is that it allows for the identification of the most important pathogen reservoirs, assisting risk managers to prioritize interventions and focus control strategies at the animal production level. Particularly, if repeated on a regular basis, the approach is regarded as a powerful tool to monitor the progress of control and follow the trends in the sources of human infections (Hald et al., 2004; Pires et al., 2009).

The results of this type of model can also provide estimates for the effect on the number of human cases originating from a particular reservoir (e.g. turkeys), if the observed prevalence in that reservoir is changed for instance following the implementation of a control program. Given the nature of the model, it will also be able to provide estimates on the expected change in human cases for specific subtypes, e.g. specific serovars of *Salmonella*.

It should be stressed that the model attributes human cases to the reservoir level meaning that the model is not able to differentiate between transmission routes within the same reservoir. For instance human cases linked to the pig reservoir will include cases both infected through consumption of pork and cases infected through direct contact with pigs. Therefore, and in contrast to a “traditional” farm-to-consumption risk assessment model, the model does not give detailed insight into transmission routes and cannot provide estimates for the expected changes in human infections by the introduction of specific intervention strategies. The model can, however, still employ data from carcass or food

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

sampling as long as the subtype distribution obtained from this sampling is assessed to reflect the subtypes distribution in the original reservoir e.g. cross-contamination from other sources is minimal.

The Hald model described above was initially designed with two dimensions: *Salmonella* subtype and food source. In 2010, the model was extended by Pires and Hald (2010) to include a temporal dimension (year) for trend analyses within a single country. By including a temporal dimension, the model was able to produce more robust results and it was assessed that even with only serotyping data available, the model could still produce meaningful results. This was considered to be useful for countries that use only serotyping in their national surveillance of *Salmonella*.

Through the EFSA service contracts CT/EFSA/BIOHAZ/2010/02 and CT/EFSA/ZOONOSES/2010/02, the Hald model was adapted to the EU level by including MS as a third dimension. The model produces attribution estimates at the overall EU level as well as MS-specific estimates, and allows for exploring the *Salmonella* contribution from food traded between MSs by accounting for export and import figures for the included food sources.

2. Development of a user-friendly interface for the Source-Attribution Model – EFSA_SAM

As described above, two mathematical models for *Salmonella* source-attribution at the EU level have been developed through two independent EFSA service contracts (the BT-SAM model from the CT/EFSA/BIOHAZ/2010/02 and the EU-*Salmonella* Source Attribution (EU-SSA) model from the CT/EFSA/ZOONOSES/2010/02). However, the two models are in principle structured in the same way and employing the same type of data.

The user-friendly model will in the following be referred to as the EFSA Source Attribution Model: EFSA_SAM. A user manual for the EFSA_SAM interface is provided in Appendix A.

2.1. Model dimensions

Both the BT-SAM and the EU-SSA are 3-dimensional models with the following dimensions:

- EU Member State (MS)
- Animal-food source
- *Salmonella* serotype

However, the two models include different numbers of MSs and *Salmonella* serotypes i.e. the length of the arrays varies. For future models it is expected that different array lengths would be needed for instance by the inclusion of more (or fewer) MSs or animal-food sources.

The EFSA_SAM model is therefore made flexible, so that a user-specified number of MSs, food-sources and *Salmonella* serotypes can be included.

There is a specific model script addressing the situation for a model with only a single MS included, as it is expected that some MSs would like to apply the model to national data only. In this single-MS model it is possible to include time (e.g. year) as a new third dimension replacing the MS dimension from the EFSA_SAM model. The single-MS model may be useful for MSs that want to apply the model on a regular basis to evaluate the effect of current *Salmonella* control.

The number of food sources that can be included is in theory unlimited, but a minimum of three food sources should be included in order to produce sensible results. The included food sources should represent the most important sources of *Salmonella* in the MSs included in the model and will for the

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

majority of countries in EU include pigs/pork meat, broilers/chicken meat and layers/eggs. The EFSA_SAM model is by default specifying pigs/pork meat, broilers/chicken meat, layers/eggs and turkeys/turkey meat, but with the possibility to add additional sources if relevant data are available.

The number of *Salmonella* subtypes that can be included is also unlimited. It will to a large extend be the data, including data availability and quality, that in the end will define the number included for a specific model. The EFSA_SAM model is developed to address *Salmonella* source attribution including the specification of *Salmonella* serovars and phage types for *S.* Enteritidis and *S.* Typhimurium. In addition, it is possible to include further subtyping levels such as antimicrobial resistance patterns and genotypes (e.g. MLVA).

The EFSA_SAM model is in theory able to address other food-borne pathogens than *Salmonella* by specifying other pathogen subtypes in the "Genus" and "Species" tabs (see user manual in Appendix A). It is emphasized, however, that there are some biological requirements (e.g. clonal dissemination) and data needs (e.g. sample representativeness of the included sources) that have to be fulfilled in order for the model to produce meaningful results for other pathogens (Pires et al., 2009). Other pathogen models based on the pathogen subtyping principles initially described for *Salmonella* by Hald et al. (2004) include *Campylobacter* source attribution in New Zealand using MLST typing (Müllner et al., 2009) and *Listeria monocytogenes* source attribution in UK using serotyping and AFLP typing (Little et al., 2010). For other pathogens, the approach still needs to be validated through further research. So even though the EFSA_SAM model will not require in depth knowledge of the syntax and technical specifications of the model code, it does require good understanding of the model principles and model limitations in order to interpret the results correctly.

Users of the interface are, therefore, recommended to read this report before starting using the interface to become familiar with the model principles and the mathematics behind, which is required in order to interpret the model results and assess the validity of the model.

2.2. The mathematics of the EFSA_SAM model

The EFSA_SAM model is set up in a Bayesian framework and estimates the number of human sporadic and domestic cases attributed to each source per country (λ_{cji}), assuming that the observed number of sporadic cases per subtype per country (o_{ci}) is Poisson distributed:

Poisson (o_{ci}) =$\sum \lambda_{ci}$, and

$$(1)\ \lambda_{ckji} = p_{kij} * M_{ckj} * a_{cj} * q_i$$

where λ_{ckji} is the expected number of human cases per subtype *i* and source *j* reported in country *c* and caused by food produced in country *k*, p_{kij} is the prevalence of subtype *i* in source *j* in country *k*, M_{ckj} is the amount of source *j* available for consumption in country *c* produced in country *k*, a_{cj} is the source-dependent factor for source *j* in country *c*, and q_i is the subtype-dependent factor for subtype *i*. When *c* is equal to *k* the food originates from the country in which the case is reported. Specification of the priors for a_{cj} and q_i is described in more details in section 2.5.

For the single-MS model, time period (most likely year) is replacing MS as a third dimension giving the following equations:

Poisson (o_{ti}) =$\sum \lambda_{ti}$, and

$$(2)\ \lambda_{tji} = p_{tij} * M_{tj} * a_{tj} * q_i$$

where *t* represents the time period.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

It should be noted that food sources in the single MS model can include imported food products, if data for these exist.

2.3. Model software for the interface

The BT-SAM, the EU-SSA and the TT-SAM models are all compatible with WinBUGS 1.4.3., which is a shareware that can be downloaded free of charge from http://www.mrc-bsu.cam.ac.uk/bugs/. WinBUGS 1.4.3. is by experience of the authors of this report more stable than the OpenBUGS version for this type of complex models and is therefore used for executing the model in this project.

The WinBUGS programming language is, however, quite complicated; particular for a 3-dimensional model as the EFSA_SAM model. The objective of this project was therefore to develop a flexible and user-friendly interface that can be used by food and/or public health scientists or food-safety risk managers with no or only little knowledge of the WinBUGS syntax.

The programming language (development environment) used for developing the user-friendly interface is Embarcaderos Delphi XE2 Enterprise. Delphi is a native Windows development tool and therefore has the advantage of very fast code execution and a very small footprint with no dependencies to external frameworks. This is in contrast to for instance languages like C# and VB.NET, where the resource consuming .NET framework is needed on the clients' computers, or the Java-language where the Java Runtime Environment also must be installed on the clients' computers. The independency to external frameworks means that the hardware requirements for applications developed in Delphi are smaller, and the risk of framework or runtime version inconsistencies is eliminated. Software produced in Delphi will work on both 32-bit and 64-bit Microsoft Windows versions. Delphi has a very long history of backwards code compatibility, ensuring excellent maintainability of existing code. Delphi was originally manufactured by Borland, and the first version appeared in 1995, and the latest (15th version known as XE2) was released in September 2011.

The developed interface generates a WinBUGS code based on the users selections and export this code with corresponding data to WinBUGS. There the model code is executed automatically inside WinBUGS. The model results are then imported from WinBUGS to the interface software for tabulation and graphical display. This approach ensures consistency in both model and data setup, eliminating the need for user knowledge of WinBUGS syntax, which can be error prone.

All model settings (i.e. selection of subtypes, MSs, food sources, etc.), the generated WinBUGS code and the data are stored in a scalable SQL database. This ensures track-keeping of all models executed through the user-friendly interface, making it possible to retrieve the exact setup at a later point in time for documentation. Also, easy-to-use backup and restore routines are implemented.

The interface software are delivered as a program file that can be installed from any kind of mass storage media (e.g. CD, USB keys) or be downloaded from an internet link to the users' personal computers.

2.4. Data input to the EFSA_SAM model

Below is a general description of the types of data that it is possible to employ in the EFSA_SAM model for source attribution analyses.

2.4.1. Reported human cases of *Salmonella* (or other pathogens)

Data on reported human cases of *Salmonella* (or other pathogens) with the following levels of details are required as input for the EFSA_SAM model (see data template in Appendix B1 for EU model and Appendix C1 for single MS-model):

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

- number of reported cases per subtype and MS (or year for single MS-model) in the study period

- number of travel-related cases per subtype and MS in the study period. For the single MS-model, travel data should be specified as: yes to travel, no to travel or unknown travel history (Appendix C1).

- number of outbreak-related per subtype and MS (or year for single MS-model) in the study period

Based on these data, EFSA_SAM will estimate the number of domestic and sporadic cases, which will then through the model be allocated to specific food sources or to a group of unknown, for instance if the subtype seen in humans is not found in any of the food sources included.

2.4.1.1. Genus, species and subtype specification

There is the possibility to include three levels of subtypes, which by default for *Salmonella* is called Serovar, Phage type and Genotype. In the EFSA_SAM, a list of *Salmonella* serovars and phage types are already included in the database and imported data will be compared to this list as explained in the user manual. For genotypes (e.g. MLST or MLVA types), the user specifies the format solely. In fact, the “Genotype” option doesn’t need to be a genotype, but could also be a third phenotypic subtyping level such as antimicrobial resistant patterns.

2.4.2. Underreporting factors

To take account for differences in human case underreporting between MSs, the EFSA_SAM model includes the possibility to employ underreporting factors for the MSs. The underreporting factors included in the EFSA_SAM are those used for the model answering the turkey target mandate (the TT-SAM model) as described in Hald et al. (2012). The methodologies used for estimating the underreporting factors are described in Havelaar et al. (in press). The underreporting factors are based on data from 2009 and included as probability distributions in the EFSA_SAM model in order to account for uncertainty around the data. A lognormal distribution was found to provide a good fit of the data and the estimated means and standard deviations were used as model input (Table 1). It is possible to revise these if/when better or more recent data become available in the future.

Underreporting factors (u_c) may differ depending on the pathogen in question. Alternative underreporting factors should therefore be considered by users, if the model is applied for other pathogens.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Table 1: Estimated means and standard deviations for the underreporting factors (u_c) applied in the EFSA_SAM model.

	Mean(Ln)	Sdev(Ln)	Mean dist	Mean data
Austria	2.1	0.8	11.2	11
Belgium	0.9	0.9	3.6	3.5
Bulgaria	6.3	0.8	734.8	718.4
Cyprus	4.9	0.8	177.2	173.3
Czech Republic	3.1	0.8	29.6	28.9
Denmark	1.2	0.8	4.5	4.4
Estonia	2.5	0.8	17.4	16.9
Finland	-1.3	0.8	0.4	0.4
France	3	0.8	27.5	26.9
Germany	2	0.8	10	9.8
Greece	6.8	0.8	1257	1229
Hungary	3.9	0.8	68.3	66.8
Ireland	1.1	1.1	5.6	5.4
Italy	4	0.8	73.4	71.8
Latvia	3.5	0.8	45.4	44.3
Lithuania	3.8	0.8	60.5	59.1
Luxembourg	1	1	4.6	4.4
Malta	5.1	0.8	227.8	222.6
Poland	4.5	0.8	116.6	114
Portugal	7.4	0.8	2131.2	2083.8
Romania	5.5	0.8	358.4	350.2
Slovakia	3.7	0.8	54.3	53.1
Slovenia	3.4	0.9	41.7	40.5
Spain	5.1	0.8	219.1	214.2
Sweden	-1	0.8	0.5	0.5
The Netherlands	3	0.8	26.8	26.2
United Kingdom	1.7	0.8	7.5	7.3

2.4.3. Prevalence data on *Salmonella* (or other pathogens)

For each specified food source included in the EFSA_SAM model, the number of units tested (i.e. sample size) and the number of positive units per subtype and MS (or year for single MS-model) are required as input (see data templates in Appendix B2 and C2). The EFSA_SAM will based on these values calculate a subtype-specific prevalences in each food source.

2.4.4. Production and trade data

Ideally, the EFSA_SAM model should employ consumption data of the specified food sources. However, national consumption data do not generally include information of the origin of the food (i.e. the country in which the food was produced), which is considered to be an essential part of the model because of the extensive trade of foods between MSs. Therefore, an approximation is recommended, where the amount available for consumption produced in a MS is estimated as:

Amount available for consumption = production – export

The amount of food imported to one MS from another MS should be estimated in order to consider trade between MSs. In the EFSA_SAM model, the prevalences of the specific subtypes are weighted by “production” and “import” figures as explained in section 2.2.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

The data specifications requires that the users of the EFSA_SAM model applies EUROSTAT data (or another relevant data source) on production, export and import to estimate the amount available for consumption in each MS by the MS of origin, which are used as input data for the model (see data template in Appendix B3). The single MS model also requires data on the food produced (see template in Appendix C3). If only domestic food sources are included in the model, national statistic on production or consumption may be used.

Technically, the EFSA_SAM model could make the estimations of foods available for consumption in a MS based on “raw” data input on production, import and export. However, during the development of the BT-SAM model, it was realised that there is often disagreement between import and export data, i.e. the amount reported as exported in one MS to another MS do not correspond to the amount reported as imported in the receiving MS. Therefore, some subjective judgement is needed when analysing the trade data and it is considered best that the decision of the final data to be included is taken by the model users.

The EFSA_SAM includes a demo dataset for production and trade data. This dataset is equivalent to the data used in the TT-SAM model (Hald et al., 2012). In this dataset, data on production of the animal-food sources were extracted by EFSA from the EUROSTAT and provided as Excel files. Production data for broilers and turkeys were taken from the 2010 AVEC report (AVEC, 2011), as the EUROSTAT data does not provide information for the separate poultry species. For pigs, the weight of slaughtered carcasses per MSs in 2010 was used as a measure of domestic production. Finally for eggs, data on the production of shell eggs were extracted from FAOSTAT[4], since these data were missing from many MSs in the EUROSTAT data.

All information on trade between MSs was extracted from EUROSTAT database[5] (dataset name: DS-016890-EU27 Trade Since 1988 By CN8). Export data as reported by the MSs were used for both estimating import and export. This was done in order to use only one table realising that there was a high degree of disagreement of the data reported in the export and import tables for each food source. The unit used for expressing the amount of food produced and exported was tonnes. The demo model does not include data on food imported from outside of EU due to lack of data on both imported amounts and *Salmonella* prevalences.

The amount available for consumption produced in a MS was as mentioned above estimated as: Production – export. In some instances, this resulted in negative production values i.e. the amount exported were larger than the amount produced within the country. In order to ensure that MSs would still have nationally produced food available in their own country, it was assumed that imported products could also be re-exported. Data availability and data used are described in Hald et al. (2012).

2.5. Specification of prior distributions, initial values and model execution

Every variable (or parameter) included as a probability distribution in a Bayesian framework needs to have a prior distribution specified in order for the model to estimate a posterior distribution.

As explained in section 1, the mathematical model behind the EFSA_SAM interface relies on the estimation of a posterior distribution for the food source (a) and subtype (q)-dependent factors, as these along with the inputted data is used for estimating the number of human cases per food source, subtype, and country. The prior distribution for a and q in the EFSA-SAM model are as default included as Uniform distribution with values ranging from 0-100.

[4] FAOSTAT data extracted 3. January 2010: http://faostat.fao.org/site/569/DesktopDefault.aspx?PageID=569#ancor

[5] http://epp.eurostat.ec.europa.eu/newxtweb/

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Remember that the *a*'s are based on two dimensions, meaning that the number of priors specified for *a* is the number of food sources multiplied with the number of countries (or years for single MS-model) included in the model.

For the subtype-dependent factor, the *q*-value for first subtype specified (e.g. Enteritidis) is by default set to 1 meaning that all the other subtypes are estimated relative to this. Since the priors are simply multiplication factors, it makes sense to select a common subtype as an anchor point. Often the datasets of the reported human cases and the food source prevalence data will also be constructed so that they include a group of "other subtypes". This will by default be specified as the last subtype in the dataset and data referred to this group is not used for estimating the attribution values meaning that human cases belonging to this group will by default be allocated to the group of "unknown". Because the *q*-value for the "other subtypes" is not included in the actual calculations, the number of priors for *q* ends up to equal the number of subtypes included minus two (i.e. minus a prior for the first and last subtype which are by default set to 1).

Finally, the underreporting factors are included as distributions in the model; one for each country. As explained in section 2.4.2. these are included as lognormal distributions and prior distributions ranging from -2 to 5 are included as default in EFSA_SAM.

For each parameter included as a distribution with specified priors, we also need to specify proper initial values. Initial values tell the model where in the specified distributions to start drawing random numbers. So the initial values should lie within the ranges for the prior distributions. Specifying appropriate initial values can be a bit tricky and often when a WinBUGS model stops running and returns an error message, the reason is that the initial values chosen resulted in implausible results given the distributions defined and/or the other calculations done in the model. Typically, if the model during an iteration results in a value of zero for some calculated parameter, this causes problems for the following estimations and then the model stops running. In the EFSA_SAM, initial values are generated automatically, but users may experience that the WinBUGS log file returns an error massage which will require a that a new set of initial values are generated as explained in the user manual (Appendix A).

When the prior distributions and the associated initial values have been specified, the user is requested to specify the number of burn-ins, iterations, and Markov chains that the model should run. Burn-ins is the number of iterations used for the model to arrive at some steady state, after which the model is producing stable results. The number of iterations is the set of iterations on which the final results will be calculated. Finally, the number of chains specifies how many independent runs of the model should be done. It is important to note that for each chain a new set of initial values are to be generated and it is emphasised that the initial values should be widely dispersed in the prior distribution model. The results of the different chains are used to monitor convergence as explained below.

Typically, for this kind of model, a burn-in of 5-10,000 iteration, 20-30,000 model iterations and 3 to 5 chains are appropriate.

After having specified the burn-ins, the number of iterations and the number of Markov chains, the model is ready for execution in WinBUGS. EFSA_SAM automatically sends the code generated on the basis of the users selections to the WinBUGS environment. Here a Markov Chain Monte Carlo (MCMC) simulation is applied to arrive at the posterior distributions for a_{cj} , q_{i}. and $u_{c.}$, and consequently the overall results of the model.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

2.6. Model check: Monitoring of convergence and goodness of fit

A first step to check how the model is performing is to examine the Kernel density plot of the posterior distribution of a_{cj} , q_i and u_c. These should depict distributions that at the left side end in a smooth curve approximating zero (Figure 1a). If the density distribution is cut off to the right the ranges of the prior distribution appear inappropriate and need to be extended (Figure 1b). This is easily done in the EFSA_SAM interface as explained in the user manual (Appendix A).

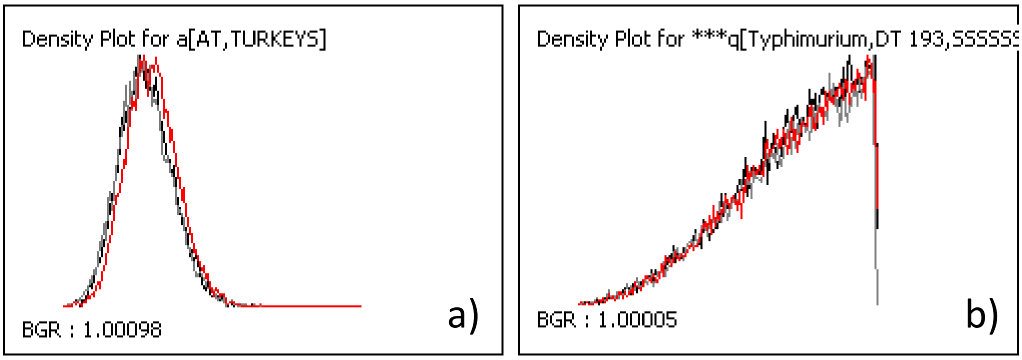


Figure 1: Kernel density plot of the posterior distribution for a) the food-source dependent factor *a* for turkeys in Austria, where the ranges of the prior distribution appears to be properly specified (plot from EFSA_SAM model), and b) the subtype dependent factor *q* for Typhimurium DT 193, fully susceptible, where the ranges of the prior distribution needs to be expanded to obtain a well-defined posterior distribution for this *q* values (plot from single MS-model).

Sometimes, and particularly in the single MS model, the density plot of a posterior distribution can appear similar to the defined prior distribution (Figure 2). This means that the model cannot find a proper value for the specific parameter (e.g. *q*). Most often this is caused by the fact that the subtype in question is found in only one single food-animal source and because the *a* and *q* values are interlinked in a multiparameter prior, the uncertainty around q in such a situation is included in the estimate for *a* -– i.e. the model doesn't care what value *q* takes. If the user can verify by scrutinizing the data, that this seems to be the case, the model results can still be considered valid (conditioning that that convergence otherwise appears to have occurred).

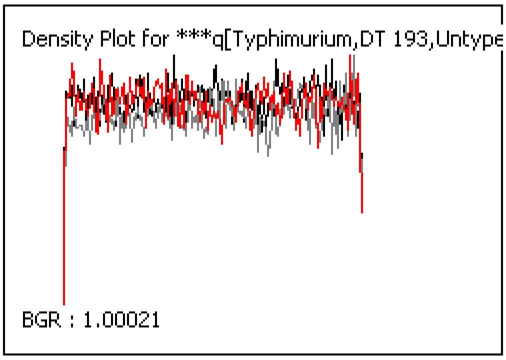


Figure 2: Kernel density plot of the posterior distribution for the subtype dependent factor *q* for Typhimurium DT193. The posterior distribution is similar to the defined prior distribution, which may indicate that this specific subtype is only found in a single food-animal source.

A crucial part of every Bayesian model is to monitor convergence i.e. to check that the model is producing robust and stable results. The use of more Markov chains and monitoring that these chains after a specified number of iterations are ending up at the same results, is an appropriate way of doing

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

this. For some modeling problems, more solutions may be available, which is why it is important to check that this is not the case - or if it is - try to explain the possible reasons.

A very simple way of monitoring convergence is by visually looking at the Kernel density plot for the a_{cj} and q_i values for each chain. If these appear to be overlying, the model seems to have converged well. In Figure 1, all three chains (indicated by different colours) overlap in an appropriate way.

A more formal approach is to calculate the bgr-diagnostics, as described by Brooks and Gelman (1998). Here convergence is considered to have occurred when the variance across all chains (B) is no larger than the variance within each individual chain (W), and when the chains had reached a stable level. In WinBUGS, the ratio R=B/W is plotted over time-iterations and will tend to 1 as convergence is approached. However, these plots cannot be exported from WinBUGS. Consequently, EFSA_SAM makes its own calculation of the bgr-diagnostic based on the results of *a* and *q* for the second half of the iterations resulting in a single value for R for each posterior distribution. The bgr-diagnostic value is presented in each density plot (Figure 1). Convergence may be assumed for practical purposes if R < 1.05 (WinBUGS user manual).

Finally, the predictive ability – or goodness of fit - of the model can be assessed by estimating the ratio between the observed human cases and the number of cases predicted by the model. A ratio around 1 indicates a good fit. A plot of the ratios from each country (or year and subtype in single MS model) can indicate if the model for some countries (or years) is predicting poorly as compared to the observed data. Poor fit is often related to data quality and the user may, therefore, want to exclude specific countries (or years) in order to evaluate their influence on the overall model results.

2.7. Baseline and scenario analyses

The data imported into EFSA_SAM will be used for a baseline analysis providing attribution estimates based on the data applied in the model. Baseline results, therefore, give an indication of the most important sources in the current situation.

The results of the baseline analysis can then be compared with the results from different scenario analyses specified by the user. The interface allows for two types of scenarios:

1. Setting target prevalences for individual subtypes. In this type of scenario, the user can change the original prevalence of specific subtypes to see the effect on the number of human cases if the prevalence for instance is reduced due to targeted control. The prevalence can be change for several subtypes in the same scenario and the new target prevalences can differ between the subtypes. The EFSA_SAM will automatically change the original prevalence to the set target prevalence, but only if the original prevalence is greater than the target prevalence. In case, the original prevalence is lower than the target prevalence, the original prevalence is kept.

2. Setting a combined target prevalence for a group of subtypes. Here the users can select any number of subtypes for which a combined prevalence should be equal or less to a set target prevalence. The EFSA_SAM generates a new set of subtype-specific prevalences that are proportionally scaled down from the original prevalences in order to result in an overall prevalence equal to or less than the target prevalence.

In both types of scenarios, the original and adjusted prevalences will be presented for the user for comparison.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

2.8. Model output and presentation of the results

Before running the model, the user is requested to select for which outcome variables statistics should be reported (Table 2 and Table 3). Statistics reported per outcome variables are mean, standard deviation, median and 95% credibility interval. Results are reported in tables that can be exported as comma-separated files for further analysis or graphical display in other softwares e.g. MS Excel.

Table 2: Description of the variables that the user is able to select as model outputs in the EU model.

Type of analysis	Scope	Variable name	Description
Baseline	EU	serotype	Number of estimated human cases in EU per source and subtype
Scenario	EU	serotypescen1	Number of estimated human cases in EU per source and subtype
Baseline	MS	source	Number of estimated human cases per source and reporting MS
Scenario	MS	sourcescen1	Number of estimated human cases per source and reporting MS
Baseline	MS	source2	Number of estimated human cases per source and MS of origin
Scenario	MS	source2scen1	Number of estimated human cases per source and MS of origin
Baseline	MS	truecasescji	Number of human cases pr. MS, food source and subtype
Scenario	MS	truecasescen1cji	Number of human cases pr. MS, food source and subtype
Baseline+scenario	MS	travel	Number of estimated travel-related human cases per MS
Baseline+scenario	MS	unknown	Number of estimated human cases with unknown source per MS
Baseline	MS	total	Number of estimated total human cases per MS
Scenario	MS	totalscen1	Number of estimated total human cases per MS
Baseline+scenario	MS	GoodFit	Goodness of fit ratio: reported number of human cases per MS divided by the estimated number of cases per MS
Baseline	EU	totalEU	Number of estimated total human cases in EU
Baseline+scenario	EU	unknownEU	Number of estimated human cases with unknown source in EU
Baseline+scenario	EU	travelEU	Number of estimated travel-related human cases in EU
Baseline+scenario	EU	unktravEU	Sum of unknownEU and travel EU
Scenario	EU	totalEUscen1	Number of estimated total human cases in EU
Scenario	EU	difftotalEU	totalEU – totalEUscen1
Baseline	EU	sourceC	Number of estimated human cases per source in EU
Baseline	EU	percsourceC	Percentage of estimated human cases per source in EU
Scenario	EU	sourceCscen1	Number of estimated human cases per source in EU
Scenario	EU	percsourceCscen1	Percentage of estimated human cases per source in EU
Scenario	EU	diffsourceC	sourceC – sourceCscen1
Scenario	EU	diffpercsourceC	percsourceC – percsourceCscen1
Scenario	EU	percdiffsourceC	Percentage difference between baseline and scenario: = diffsourceC*100/source
Baseline+scenario	EU+MS	a	Food-source related factor per source and MS
Baseline+scenario	EU+MS	q	Subtype related factor; one per subtype. The q-value for the first subtype (in the demo data *S.* Enteritidis) is by default set to 1 and the others q-values are estimated relative to this one
Baseline+scenario	EU+MS	u	Underreporting factor; one per MS

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Table 3: Description of the variables that the user is able to select as model outputs in the single MS model.

Variable name	Description
source	Number of estimated human cases per source per year
travel	Number of estimated travel-related human cases per year
outbreak	Number of estimated outbreak-related human cases per year
unknown	Number of estimated human cases with unknown source per year
total	Number of estimated total human cases per year
percsource	Percentage of estimated human cases per source per year
lambdatji	Number of estimated sporadic and domestic human cases per subtype, source and year
lambdaexp	Number of estimated human cases per subtype and year
a	Food-source related factor; one per source
q	Subtype related factor; one per subtype. The q-value for the first subtype (in the demo data *S*. Enteritidis) is by default set to 1 and the others q-values are estimated relative to this one

2.9. Saving and exporting an analysis

After having finished an analysis, the whole analysis is automatically saved. The analysis including set up and results can also be exported as individual files in an easy to use format for further analysis or graphical display in other softwares e.g. Excel. The whole analysis including results can also be saved as a zip file for future documentation purposes.

Conclusions and Recommendations

The source attribution approach based on subtyping has been assessed to be a valuable tool for pointing out the most important source of human salmonellosis in the EU as well as to predict changes in human incidence, if prevalences in different animals reservoirs and for specific serovars are changed for instance as a result of targeted control. The principle has also been applied in individual countries for prioritizing risk management strategies.

A drawback of the models used for source attribution is that these models require specialist knowledge both in statistics and programming. However, with the development of EFSA_SAM, a flexible user-friendly interface for attributing human cases of food-borne pathogens to the responsible food-animal reservoirs and/or food sources is now available, which will hopefully make the approach accessible and useful for a broader audience.

It is emphasised though, that the interpretation and validation of the model results require some statistical knowledge, but with this technical report and the user manual at hand, the users should be able to apply and run the model, as well as making an appropriate evaluation the results.

It is also stressed, that like for all models, the results are never more valid than the data that is applied. Data validation is, therefore, an important part of any source attribution exercise. Inclusion of all major animal-reservoir sources for human infections and the availability of representative data on prevalence and subtype distribution from the same reservoirs are therefore a requirement. Likewise, incidence data and subtype distribution from human cases is needed.

Finally, the usefulness of the model has also so far only truly been evaluated for *Salmonella*. Application of the model for other pathogens, therefore, needs further research and evaluation before the results can be used for risk management decision support.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

The results of EFSA_SAM are expected to be useful for the delineation of risk management strategies, particularly if the model is applied on a regular basis, to evaluate the impact of targeted interventions and dynamic changes in the sources of human salmonellosis.

It is recommended that all user of EFSA_SAM read both this technical report and the user manual before starting to use the software.

References

AVEC (Association of Poultry Processors and Poultry Trade in the EU Countries), 2011. 2011 Annual Report. Available from http://www.avec-poultry.eu/Default.aspx?ID=4731, 52 pp.

Brooks SP and Gelman A, 1998. Alternative methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics, 7, 434-455.

Hald T, Vose D, Wegener HC and Koupeev T, 2004. A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. Risk Analysis, 24, 255-269.

Hald T, Pires SM and de Knegt L, 2012. Development of a *Salmonella* source-attribution model for evaluating targets in the turkey meat production. Report to contract CT/EFSA/BIOHAZ/2011/02. Supporting publications, 2012:259, 36 pp.

Havelaar AH, Ivarsson S, Löfdahl M and Nauta MJ, in press. Estimating the true incidence of campylobacteriosis and salmonellosis in the EU, 2009. Epidemiology and Infection, published online 13 April 2012. DOI: http://dx.doi.org/10.1017/S0950268812000568.

Little CL, Pires SM, Gillespie IA, Grant K and Nichols GL, 2010. Source attribution of *Listeria monocytogenes* in England and Wales: Adaptation of the Hald Salmonella Source Attribution Model. Foodborne Pathogens and Disease 7(7), 749-756.

Müllner P, Jones G, Noble A, Spencer SEF, Hathaway S, French NP, 2009. Source attribution of food-borne zoonoses in New Zealand: a modified Hald model. Risk Analysis 29(7), 970-984.

Pires SM and Hald T, 2010. Assessing the differences in public health impact of Salmonella subtypes using a Bayesian microbial subtyping approach for source attribution. Foodborne Pathogens and Disease, 7, 143-151.

Pires SM, Evers EG, van Pelt W, Ayers T, Scallan E, Angulo FJ, Havelaar A and Hald T, 2009. Attributing the human disease burden of foodborne infections to specific sources. Foodborne Pathogens and Diseases, 6, 417-424.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

APPENDICES

A. EFSA_SAM MANUAL VERSION 1.0

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

User friendly Interface for

EFSA Source Attribution Modeling

EFSA_SAM

Manual version 1.0

Software Version 1.0 - 20120623

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

TABLE OF CONTENTS

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

INTRODUCTION

The flexible user friendly interface software for the EFSA Source Attribution Model (hereafter called EFSA_SAM) is split into two parts: A *general user part* and an *advanced user part*.

The reason for this is that higher flexibility often leads to higher complexity, and higher complexity often leads to lower user friendliness. Also, very often, software in general is going to be used by two groups of people: The occasional users that just needs to perform a simple task, and the advanced users that need to be able to adjust a lot of details.

Transferring this to EFSA_SAM, the software is split into two main areas.

One (general) area is used for running analyses based on predefined models (importing data, setting up analysis dimensions and running the model).

The other (advanced) area is used for setting up the models (e.g. model code, WinBUGS Scripting, WinBUGS/EFSA_SAM interface variables etc).

It is emphasized that even if the software seems simple to use, it is required that the users know the data requirements and have some statistical background for running the analyses, and know how to identify the pitfalls and evaluate the analysis results. It is, therefore, recommended, that users also read the technical report describing the modeling principles and mathematics in more detail before starting to use the software.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

1. INSTALLATION

The user needs to have both WinBUGS and EFSA_SAM installed on the PC, as EFSA_SAM call WinBUGS for execution of the model code developed in EFSA_SAM. It does not matter whether WinBUGS or EFSA_SAM is installed first.

1.1. WinBUGS installation

Prerequisites: WinBUGS 1.4.3[6] must be installed, in order to use EFSA_SAM.

WinBUGS is a freeware and a lot of resources can be found here http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml

WinBUGS – Download and installation

Download http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/winbugs14.zip

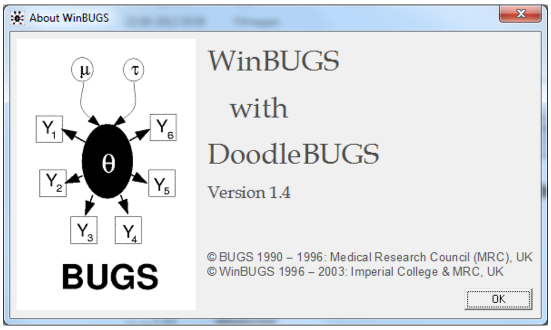
This version will work on both Windows XP, Windows Vista and Windows 7 (32 and 64-bit versions), and you will be able to control exactly where WinBUGS will be located.

Create a folder C:\EFSA_SAM\Winbugs14 and unpack the contents of the zip-file to this folder.

Start your copy of WinBUGS14 by double clicking C:\EFSA_SAM\Winbugs14\Winbugs14.exe.

Verifying the WinBUGS version

When WinBUGS is running, press *"Help"* and then *"About WinBUGS"*
The version number must be 1.4.3. If you just see 1.4, then you need to upgrade WinBUGS to 1.4.3.



WinBUGS
with
DoodleBUGS
Version 1.4

© BUGS 1990 – 1996: Medical Research Council (MRC), UK
© WinBUGS 1996 – 2003: Imperial College & MRC, UK

How to upgrade

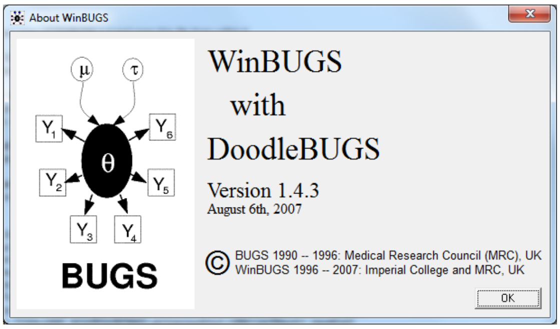
Navigate to http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/WinBUGS14_cumulative_patch_No3_06_08_07_RELEASE.txt

[6] Lunn DJ, Thomas A, Best N and Spiegelhalter D, 2000. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing, 10, 325-337.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Select everything and save it to a text file.

Then follow the instructions. After a successful upgrade you should see Version 1.4.3 in the *"About WinBUGS"* window.



WinBUGS
with
DoodleBUGS

Version 1.4.3
August 6th, 2007

© BUGS 1990 -- 1996: Medical Research Council (MRC), UK
WinBUGS 1996 -- 2007: Imperial College and MRC, UK

License key - download and installation
Navigate to http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/WinBUGS14_immortality_key.txt
Select all text and copy it to your clipboard.

Click File/New.

Paste the text into the blank window.

From the Tools menu pick the Decode option. A dialog box will appear.

Click on the "Decode All" button to install the key.

Quit and restart WinBUGS to start using the full version.

Failing to install the license key properly will result in an error message and termination of WinBUGS when trying to run a model, regardless whether it is done directly from WinBUGS or via EFSA_SAM.

1.2. EFSA_SAM installation

Insert the CD-ROM or USB-Key.

Run EFSA_SAM_Setup.exe. The rest of the installation will run automatically.

A shortcut to the program will be placed on the user's desktop.

The software will, by default, be installed in the C:\EFSA_SAM folder.

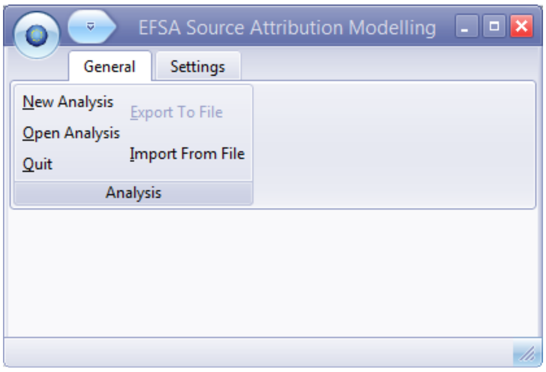
2. The user interface

When starting EFSA_SAM, a simple main window will appear. It contains a top menu containing two main items: *General* and *Settings.*

The *"General"* menu has three sub menu items:

- New Analysis - Creates a new analysis
- Open Analysis - Opens a previously created analysis
- Import From File - Will copy an already existing analysis, making it possible to change settings and rerun the analysis.

- Export To File – Exports the current analysis to a zip-file. This includes the model, data and results.
- Quit – Closes EFSA_SAM.



The *"Settings"* menu is where the more advanced features are located. That is, creating and modifying models, maintaining the more general data such as countries, food sources, bacteria types and subtypes etc. Also, database backup and restore functionality is located here.

2.1. The workflow sidebar

The *workflow sidebar* is where the user controls the work flow in an easy to understand sequence of tasks to perform.

First data imports such as number of reported human cases/incidence, food source prevalences and food import/export values.

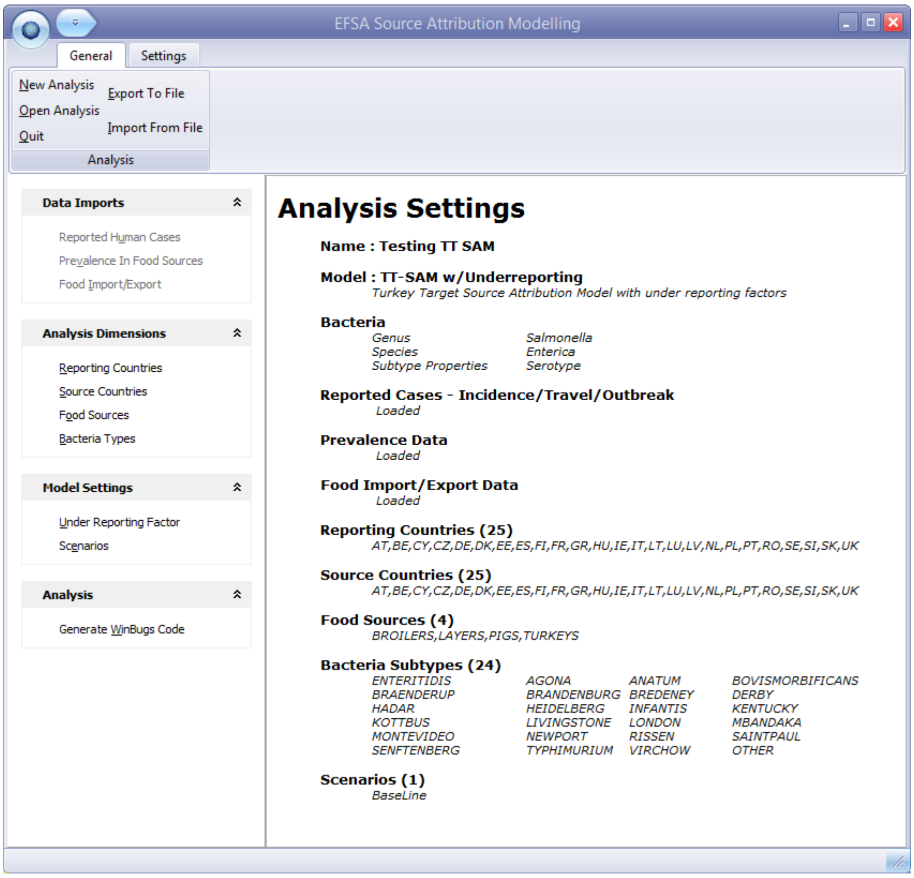
Then, if needed, changing the analysis dimensions such as reporting- and source countries, food sources and bacteria types.

Next thing to set up is the model settings such as revising the underreporting factors if these are applied (see *select model* below), and then setting up different scenarios to analyse along with the baseline analysis (see *scenarios* below).

The final step is to generate the WinBUGS code and execute the analysis in WinBUGS.

Afterwards, the WinBUGS results are imported into the user friendly interface.

The number of items present in the sidebar will change depending on the model selected. It is fully configurable when setting up the models in the advanced part of EFSA_SAM. You can read more about this later in the manual.



Analysis Settings

Name : Testing TT SAM

Model : TT-SAM w/Underreporting
Turkey Target Source Attribution Model with under reporting factors

Bacteria

Genus	*Salmonella*
Species	*Enterica*
Subtype Properties	*Serotype*

Reported Cases - Incidence/Travel/Outbreak
Loaded

Prevalence Data
Loaded

Food Import/Export Data
Loaded

Reporting Countries (25)
AT,BE,CY,CZ,DE,DK,EE,ES,FI,FR,GR,HU,IE,IT,LT,LU,LV,NL,PL,PT,RO,SE,SI,SK,UK

Source Countries (25)
AT,BE,CY,CZ,DE,DK,EE,ES,FI,FR,GR,HU,IE,IT,LT,LU,LV,NL,PL,PT,RO,SE,SI,SK,UK

Food Sources (4)
BROILERS,LAYERS,PIGS,TURKEYS

Bacteria Subtypes (24)

ENTERITIDIS	*AGONA*	*ANATUM*	*BOVISMORBIFICANS*
BRAENDERUP	*BRANDENBURG*	*BREDENEY*	*DERBY*
HADAR	*HEIDELBERG*	*INFANTIS*	*KENTUCKY*
KOTTBUS	*LIVINGSTONE*	*LONDON*	*MBANDAKA*
MONTEVIDEO	*NEWPORT*	*RISSEN*	*SAINTPAUL*
SENFTENBERG	*TYPHIMURIUM*	*VIRCHOW*	*OTHER*

Scenarios (1)
BaseLine

2.2. The center information window

The *center information window* is where the user can follow the progress of the model setup. The window information is updated every time the user makes a change in the setup, and will therefore give a good overview of the settings selected, and how far the user is in the process.

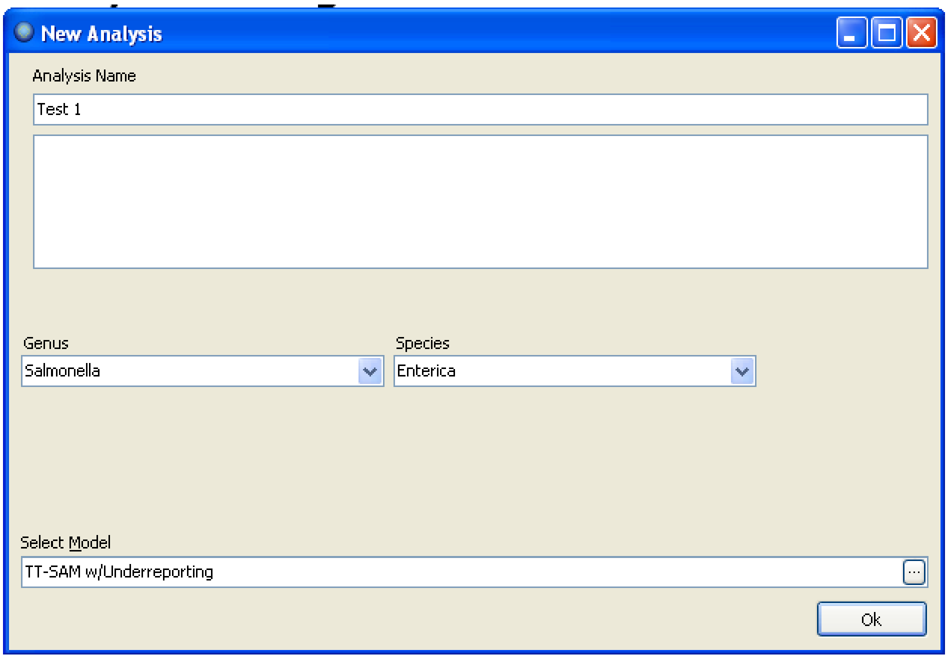
3. Creating a new analysis

Click “*New Analysis”* from the *”General”*-menu.

In the window that appears, enter a name and description for the analysis, select Genus and Species of the bacteria under study from the drop downs. (It should be noted that for most pathogens other than *Salmonella*, the approach still needs to be validated through further research – see also the technical report).

Then click in the *”Select Model”*-edit box (or press the ellipsis button to the right).

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.



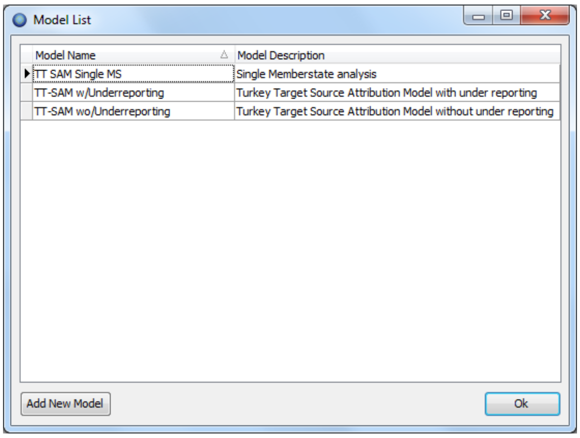
Analysis Name: Test 1

Genus: Salmonella

Species: Enterica

Select Model: TT-SAM w/Underreporting

A window with a list of models appears. Double click on the model you want to run.



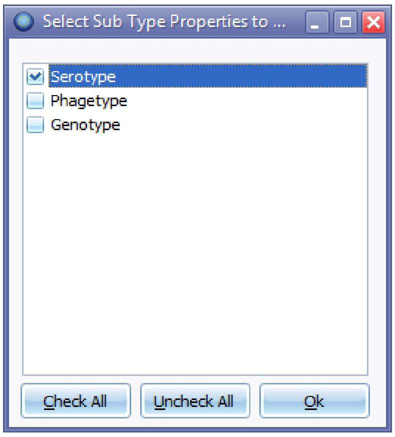
Model Name	Model Description
TT SAM Single MS	Single Memberstate analysis
TT-SAM w/Underreporting	Turkey Target Source Attribution Model with under reporting
TT-SAM wo/Underreporting	Turkey Target Source Attribution Model without under reporting

There exist three demo models in the software:

- TT-SAM Single MS: In the single member state model, the user can specify a model including data from a single country only, but for more years.
- TT-SAM w/underreporting: Here the user can specify a model for several countries accounting for underreporting of human cases.
- TT-SAM wo/underreporting: Here the user can specify a model for several countries, but choose not to account for underreporting of human cases.

After clicking the appropriate model, a new window appears. Here you need to select which properties of your bacteria subtypes you want to use. The properties you select here, must be present as data columns in the "*Reported human cases*" and *"Prevalence in food sources"*-data files that you will be importing later in the process.

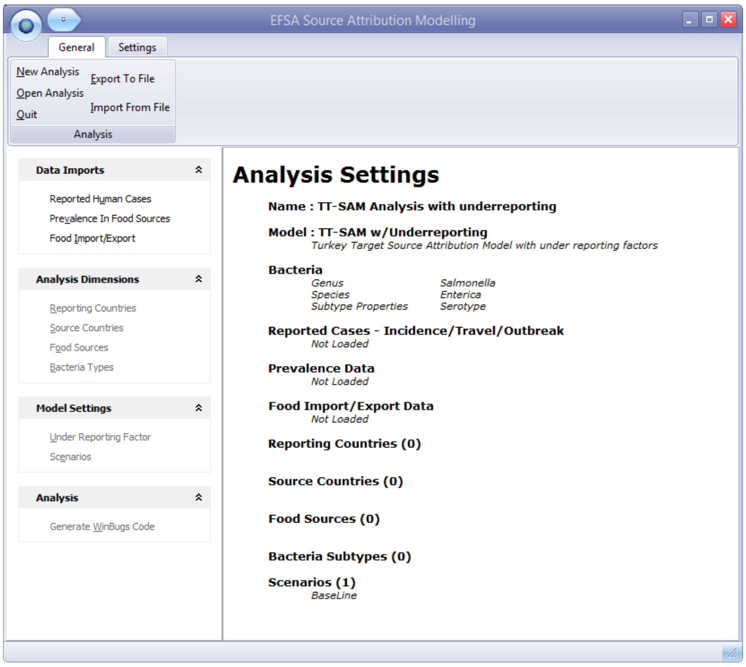
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.



☑ Serotype
□ Phagetype
□ Genotype

After pressing *"Ok"*, the window will disappear, and the sidebar and center information window will appear. At this point information about the bacteria types and the chosen model will be present.

Also, a list of settings you are going to set up will be present.



Analysis Settings

Name : TT-SAM Analysis with underreporting

Model : TT-SAM w/Underreporting
Turkey Target Source Attribution Model with under reporting factors

Bacteria

Genus	*Salmonella*
Species	*Enterica*
Subtype Properties	*Serotype*

Reported Cases - Incidence/Travel/Outbreak
Not Loaded

Prevalence Data
Not Loaded

Food Import/Export Data
Not Loaded

Reporting Countries (0)

Source Countries (0)

Food Sources (0)

Bacteria Subtypes (0)

Scenarios (1)
BaseLine

3.1. Data imports

The import windows for "*Reported human cases*", "*Prevalence in food sources*" and "*Food Import/Export*" are more or less the same, so there is no need to cover all three of them in detail.

However, all the different features of the imports will be covered in detail.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

3.1.1. Data format and validation

EFSA_SAM can import data in CSV file format. The columns needed in the files depend on the model chosen, and must be documented for each model created. The List Separator in the files must be Semicolon (;), the Decimal Separator must be Period (.) and the Thousands Separator must be Comma (,).

The first row of the data file must contain the column names (also separated by semicolons). The sequence of the columns does not matter, but all columns must be present, and preferably correctly spelled.

Datasets must be "clean" and verified before importing them into EFSA_SAM, although a number of checks will be performed to validate the data entry and headings: i) Column Naming, ii) Text values for the dimensions (Country, Subtype, Food Source), iii) Data range checks for e.g. Year, iv) Negative values, v) sums/differences of some columns, and vi) duplicate data on the combination of dimensions.

In order to ensure proper results from the analysis, we need to ensure that the terms used for countries, subtype names etc. are the same for prevalences, human cases and food import/export. For instance, if the term *S. Enteritidis* is used in the prevalences file, and the human cases file uses the term *Enteritidis*, the software will never find a match between Prevalence and Human Cases for this particular subtype.

Fortunately, EFSA_SAM makes it easier to ensure consistency, as all country names and the most frequently used subtypes (*Salmonella* serovars and phage types) and food source names are stored in a database. The data from all imported files will be checked against this database, and as you will see, EFSA_SAM can replace differently spelled items throughout the whole file in a few mouse clicks. Also new items (e.g. additional subtypes or food sources) can be added to the database for future use. Even though this is a very valuable function, it is highly recommended to use the same naming conventions throughout the datasets.

Also, if the data file contains columns named differently from the columns required, a window will ask the user to substitute the column names.

If disagreeing data entries are neither substituted nor added as new to the database, the data file cannot be imported.

Note: EFSA_SAM can handle both full country names and the country short codes (e.g. Denmark and DK will be handled the same way, and does not need substitution).

3.1.2. Importing data

In order to provide a concrete example, we will use the *Salmonella* Turkey Target model (TT-SAM) as a reference (Hald et al., 2012)[7].

The data columns needed for this model is as follows:

[7] Hald T, Pires SM and de Knegt L, 2012. Development of a Salmonella source-attribution model for evaluating targets in the turkey meat production. Report to contract CT/EFSA/BIOHAZ/2011/02. Supporting publications, 2012:259, 36 pp.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Required columns for food source prevalences

- Country (Either two-character short code or the full country name)
- Food_Source (Text - Broilers, Pigs, Turkeys, Layers; more food sources can be added by the user)
- Genus (Text - as defined for the bacteria in the software; Genus=*Salmonella* for the TT-SAM)
- Species (Text - as defined for the bacteria in the software; Species=*Enterica* for the TT-SAM)
- SubTypeFields (like Serotype, Phagetype, Genotype - as chosen when you created the analysis. For genotypes (e.g. MLST or MLVA types), the user specifies the format solely. In fact, the "Genotype" option doesn't need to be a genotype, but could also be a third phenotypic subtyping level such as antimicrobial resistant patterns. For the TT-SAM, serotypes are used. (Other bacteria types than *Salmonella* can be defined, and the subtype names of these can be defined by the user and are not included in the EFSA_SAM database - see more about this in the advanced section of the manual. It is also emphasized again that the modeling approach still needs to be validated for most other pathogens).
- Units_Tested (integer number; number of samples tested, will be the denominator for calculating the prevalence per food source and country)
- Units_Positive (integer number; number of positive samples, will be the nominator for calculating the prevalence per subtype, food source and country)

The food source prevalence in percent is calculated by EFSA_SAM and is outputted as a floating point numbers with 4 decimals.

For the Single MS model, Country is replaced by Year (four digits)

Required columns for reported human cases: Incidence/travel/outbreak

- Country (Either two-character short code or the full country name)
- Genus (Text - as defined for the bacteria in the software)
- Species (Text - as defined for the bacteria in the software)
- Subtype Fields (like Serotype,Phagetype,Genotype – same as for Prevalence data described above)
- Incidence (integer number; reported number of cases per subtype and country)
- Travel (integer number; reported number of cases related to travel abroad per subtype and country)
- Outbreak (integer number; reported number of cases related to outbreaks per subtype and country)

  For the Single member state model, Travel is replaced by Travel_Yes, Travel_No and Travel_Unknown. Also, Country is replaced by Year (4 digits)

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

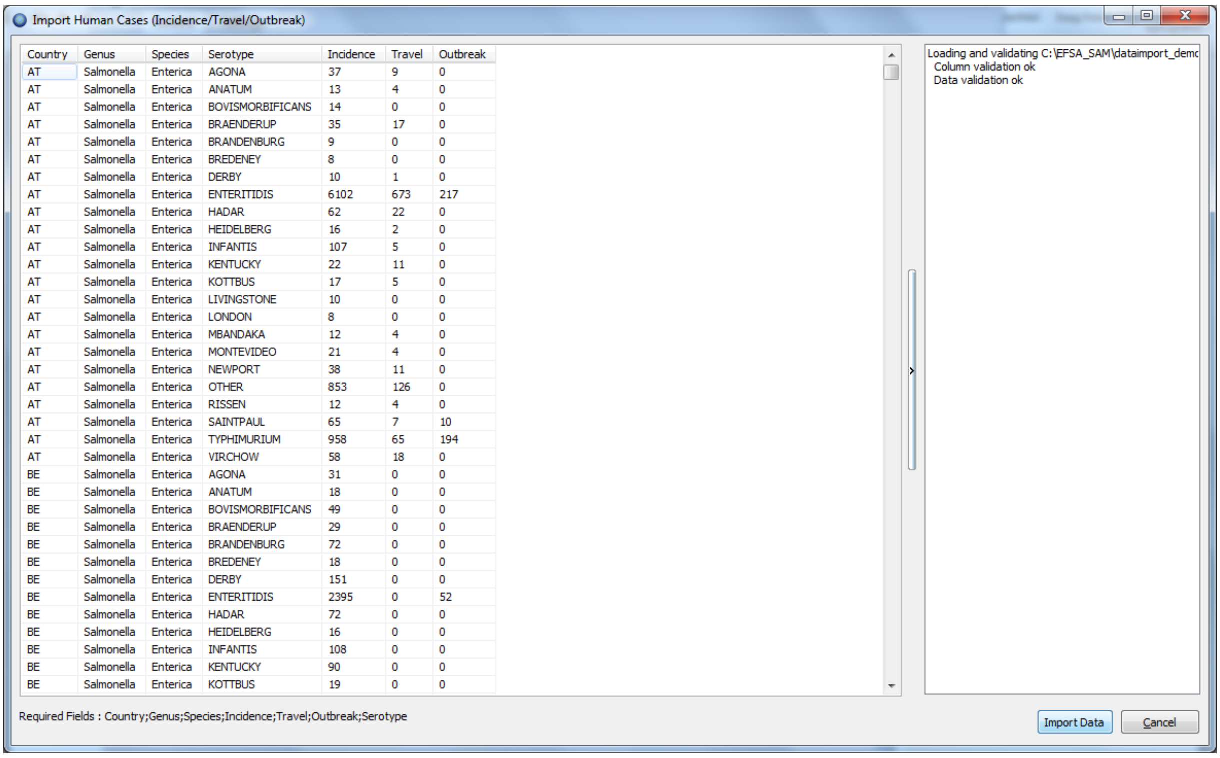
Required columns for food production and import/export data

- Country_From (Either two-character short code or the full country name)
- Country_To (Either two-character short code or the full country name)
- Food_Source (Text - Broilers, Pigs, Turkeys, Layers; more food sources can be defined by the user)
- Tonnes (Integer number - no decimals; amount in tonnes of food source)

When Country_From and Country_To is the same, the required value is the amount produced and available for consumption in the country (i.e. domestic production), typically estimated as total production minus export.

For the Single MS model Country is replaced by Year (4 digits).

When choosing one of the data imports from the sidebar, a file dialog will appear. When the appropriate file has been chosen, the data import window will appear.



Country	Genus	Species	Serotype	Incidence	Travel	Outbreak
AT	Salmonella	Enterica	AGONA	37	9	0
AT	Salmonella	Enterica	ANATUM	13	4	0
AT	Salmonella	Enterica	BOVISMORBIFICANS	14	0	0
AT	Salmonella	Enterica	BRAENDERUP	35	17	0
AT	Salmonella	Enterica	BRANDENBURG	9	0	0
AT	Salmonella	Enterica	BREDENEY	8	0	0
AT	Salmonella	Enterica	DERBY	10	1	0
AT	Salmonella	Enterica	ENTERITIDIS	6102	673	217
AT	Salmonella	Enterica	HADAR	62	22	0
AT	Salmonella	Enterica	HEIDELBERG	16	2	0
AT	Salmonella	Enterica	INFANTIS	107	5	0
AT	Salmonella	Enterica	KENTUCKY	22	11	0
AT	Salmonella	Enterica	KOTTBUS	17	5	0
AT	Salmonella	Enterica	LIVINGSTONE	10	0	0
AT	Salmonella	Enterica	LONDON	8	0	0
AT	Salmonella	Enterica	MBANDAKA	12	4	0
AT	Salmonella	Enterica	MONTEVIDEO	21	4	0
AT	Salmonella	Enterica	NEWPORT	38	11	0
AT	Salmonella	Enterica	OTHER	853	126	0
AT	Salmonella	Enterica	RISSEN	12	4	0
AT	Salmonella	Enterica	SAINTPAUL	65	7	10
AT	Salmonella	Enterica	TYPHIMURIUM	958	65	194
AT	Salmonella	Enterica	VIRCHOW	58	18	0
BE	Salmonella	Enterica	AGONA	31	0	0
BE	Salmonella	Enterica	ANATUM	18	0	0
BE	Salmonella	Enterica	BOVISMORBIFICANS	49	0	0
BE	Salmonella	Enterica	BRAENDERUP	29	0	0
BE	Salmonella	Enterica	BRANDENBURG	72	0	0
BE	Salmonella	Enterica	BREDENEY	18	0	0
BE	Salmonella	Enterica	DERBY	151	0	0
BE	Salmonella	Enterica	ENTERITIDIS	2395	0	52
BE	Salmonella	Enterica	HADAR	72	0	0
BE	Salmonella	Enterica	HEIDELBERG	16	0	0
BE	Salmonella	Enterica	INFANTIS	108	0	0
BE	Salmonella	Enterica	KENTUCKY	90	0	0
BE	Salmonella	Enterica	KOTTBUS	19	0	0

```
Loading and validating C:\EFSA_SAM\dataimport_demc
  Column validation ok
  Data validation ok
```

Required Fields : Country;Genus;Species;Incidence;Travel;Outbreak;Serotype

In the right side of the data import window, a status of the import progress can be seen.

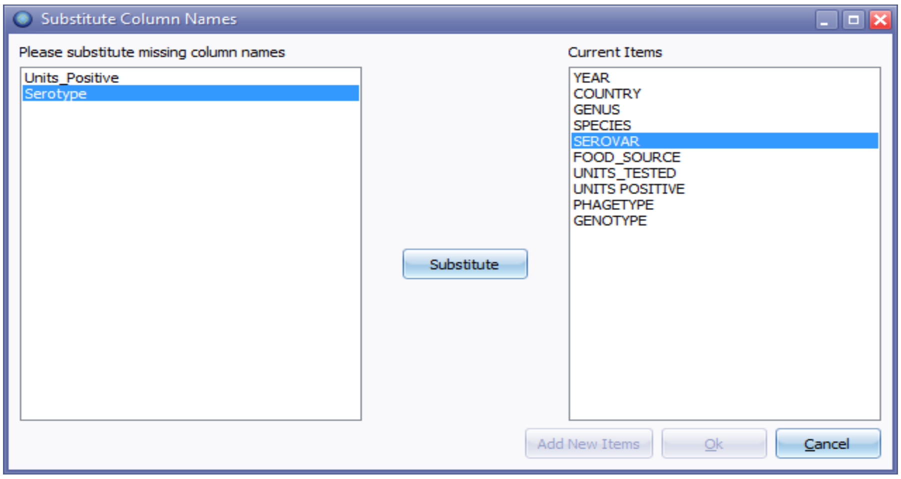
A list of the fields required can be seen at the bottom of the window. The required fields will change accordingly to the model chosen and the type of data that you are going to import.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

3.1.3. Column name substitution

After selecting a file, EFSA_SAM will check if the columns required are present in the file. If one or more columns are missing (a misspelled column will appear as missing), the user will be given the choice to substitute one or more column names in the file with the names that the software expects. This will not change anything in the file on the disk, but just serve as a way to explain the software what to look for instead.

The EFSA_SAM will display the column names it expects throughout the rest of the analysis setup (i.e. if a column in the file was called *Serovar* and a column name *Serotype* was expected, the term *Serotype* will be used).



Please substitute missing column names

Units_Positive
Serotype

Current Items

YEAR
COUNTRY
GENUS
SPECIES
SEROVAR
FOOD_SOURCE
UNITS_TESTED
UNITS POSITIVE
PHAGETYPE
GENOTYPE

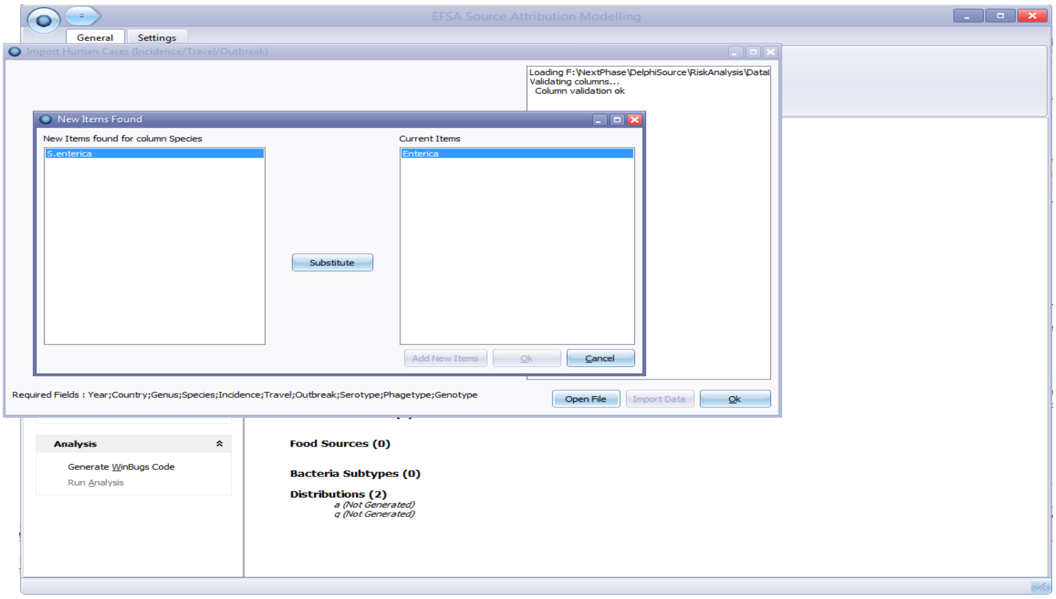
In the example shown above, you can see to the left, that two columns in the file were missing (Units_Positive and Serotype), and the columns that were found in the file. Here we will just substitute *SEROVAR* with *Serotype*, and *UNITS POSITIVE* with *Units_Positive* (note the missing underscore). The column names are case insensitive.

3.1.4. Dimension data value substitution

Just like column names can be substituted, dimension data values can also be substituted.

In the example shown below, an entry called *”S. Enterica”* has been found in the import file, and the software tells that a value *”Enterica”* exists in the database, and a substitution is possible.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.



```
Loading F:\NextPhase\DelphiSource\RiskAnalysis\Data\
Validating columns...
  Column validation ok
```

New Items found for column Species: S.enterica

Current Items: Enterica

Required Fields : Year;Country;Genus;Species;Incidence;Travel;Outbreak;Serotype;Phagetype;Genotype

Food Sources (0)

Bacteria Subtypes (0)

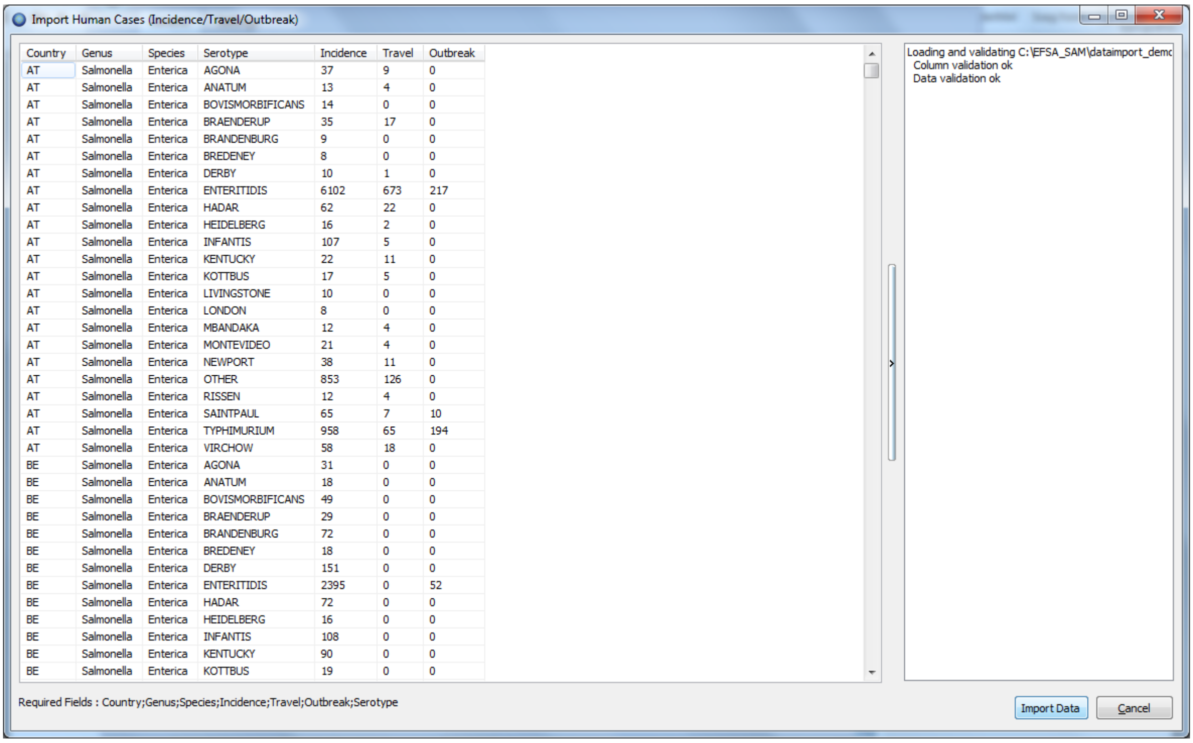
Distributions (2)
a (Not Generated)
q (Not Generated)

If an appropriate substitution is not available, it is possible to add the new entry into the database. This value is now available for all future data imports.

It is important to note that you must not add names that you can find a substitution for. This will obscure the database, and you will not be able to trust the data mapping in the future for this particular value, which in turn will produce unreliable results.

So even though EFSA_SAM does its best to ensure that data is mapped correctly, it is highly recommended to use the same naming convention for the dimension values (countries, food sources and serovars) throughout all the datasets.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.



Country	Genus	Species	Serotype	Incidence	Travel	Outbreak
AT	Salmonella	Enterica	AGONA	37	9	0
AT	Salmonella	Enterica	ANATUM	13	4	0
AT	Salmonella	Enterica	BOVISMORBIFICANS	14	0	0
AT	Salmonella	Enterica	BRAENDERUP	35	17	0
AT	Salmonella	Enterica	BRANDENBURG	9	0	0
AT	Salmonella	Enterica	BREDENEY	8	0	0
AT	Salmonella	Enterica	DERBY	10	1	0
AT	Salmonella	Enterica	ENTERITIDIS	6102	673	217
AT	Salmonella	Enterica	HADAR	62	22	0
AT	Salmonella	Enterica	HEIDELBERG	16	2	0
AT	Salmonella	Enterica	INFANTIS	107	5	0
AT	Salmonella	Enterica	KENTUCKY	22	11	0
AT	Salmonella	Enterica	KOTTBUS	17	5	0
AT	Salmonella	Enterica	LIVINGSTONE	10	0	0
AT	Salmonella	Enterica	LONDON	8	0	0
AT	Salmonella	Enterica	MBANDAKA	12	4	0
AT	Salmonella	Enterica	MONTEVIDEO	21	4	0
AT	Salmonella	Enterica	NEWPORT	38	11	0
AT	Salmonella	Enterica	OTHER	853	126	0
AT	Salmonella	Enterica	RISSEN	12	4	0
AT	Salmonella	Enterica	SAINTPAUL	65	7	10
AT	Salmonella	Enterica	TYPHIMURIUM	958	65	194
AT	Salmonella	Enterica	VIRCHOW	58	18	0
BE	Salmonella	Enterica	AGONA	31	0	0
BE	Salmonella	Enterica	ANATUM	18	0	0
BE	Salmonella	Enterica	BOVISMORBIFICANS	49	0	0
BE	Salmonella	Enterica	BRAENDERUP	29	0	0
BE	Salmonella	Enterica	BRANDENBURG	72	0	0
BE	Salmonella	Enterica	BREDENEY	18	0	0
BE	Salmonella	Enterica	DERBY	151	0	0
BE	Salmonella	Enterica	ENTERITIDIS	2395	0	52
BE	Salmonella	Enterica	HADAR	72	0	0
BE	Salmonella	Enterica	HEIDELBERG	16	0	0
BE	Salmonella	Enterica	INFANTIS	108	0	0
BE	Salmonella	Enterica	KENTUCKY	90	0	0
BE	Salmonella	Enterica	KOTTBUS	19	0	0

```
Loading and validating C:\EFSA_SAM\dataimport_demc
  Column validation ok
  Data validation ok
```

Required Fields : Country;Genus;Species;Incidence;Travel;Outbreak;Serotype

After column names and dimension values have been validated, the file is displayed and the *Import Data* button is now active. The data will not be imported before this button is pressed.

When the import button is pressed, the data will finally be loaded into the analysis, and the data import windows will disappear if the import succeeds. If the import fails, the window will not disappear, and the status window to the right, will make a note of what went wrong.

Adding values and data substitution does not change the data file on the disk.

3.1.5. Filling in the blanks

During the import procedure, EFSA_SAM will add "zero values" to non-existent data. E.g. if a prevalence for a serotype/foodsource/country is found in the "*food source prevalence*"-import file, and that same serovar/country is not found in the "*reported human cases*"-import file, then "zero values" for human cases will be added for this serovar/country.

Same thing happens if a serovar/country is present in the *"reported human cases"* data, but not in the *"food source prevalence"* data. Then "zero values" will be added for that serovar/country for all food sources.

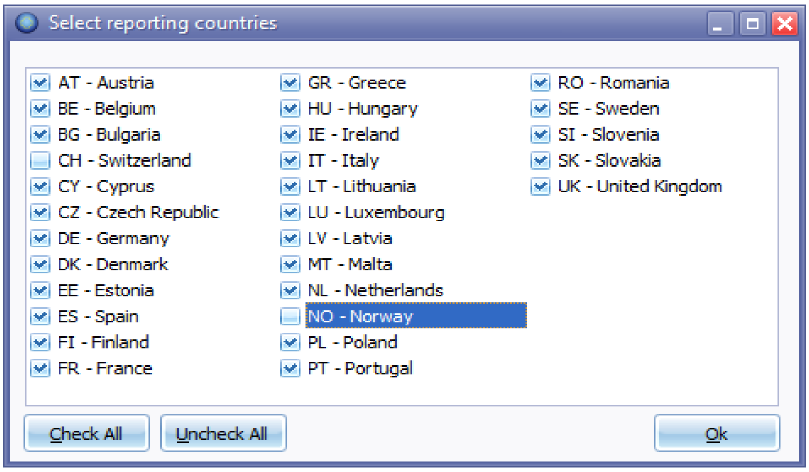
If Import/Export information is missing from/to one or more countries, these are also replaced with amounts of zero.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

4. Selection the model dimensions

4.1. Selecting reporting and source countries

The windows for selecting reporting and source countries are the same, and are pretty much self-explaining.



Select reporting countries

☑ AT - Austria
☑ BE - Belgium
☑ BG - Bulgaria
□ CH - Switzerland
☑ CY - Cyprus
☑ CZ - Czech Republic
☑ DE - Germany
☑ DK - Denmark
☑ EE - Estonia
☑ ES - Spain
☑ FI - Finland
☑ FR - France
☑ GR - Greece
☑ HU - Hungary
☑ IE - Ireland
☑ IT - Italy
☑ LT - Lithuania
☑ LU - Luxembourg
☑ LV - Latvia
☑ MT - Malta
☑ NL - Netherlands
□ NO - Norway
☑ PL - Poland
☑ PT - Portugal
☑ RO - Romania
☑ SE - Sweden
☑ SI - Slovenia
☑ SK - Slovakia
☑ UK - United Kingdom

Depending on the data input or model choice, it might be the case that it can only handle one country at a time. Hence only one country can be selected. When this is the case, the window title (as well as the title in the sidebar) will reflect this – so instead of *Select Reporting Countries* the title will be *Select Reporting Country*.

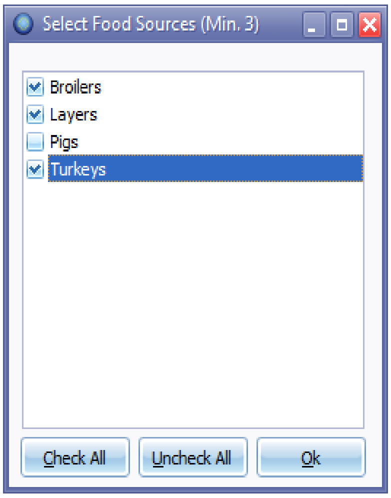
The number of countries presented will also depend on the data imported. If only 10 countries are present in the data files, then only 10 countries will be listed.

By default, EFSA_SAM automatically selects all countries present in the dataset, but it is possible to deselect countries, if needed.

4.2. Selecting the food sources

A quite self explaining window.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.



☑ Broilers
☑ Layers
□ Pigs
☑ Turkeys

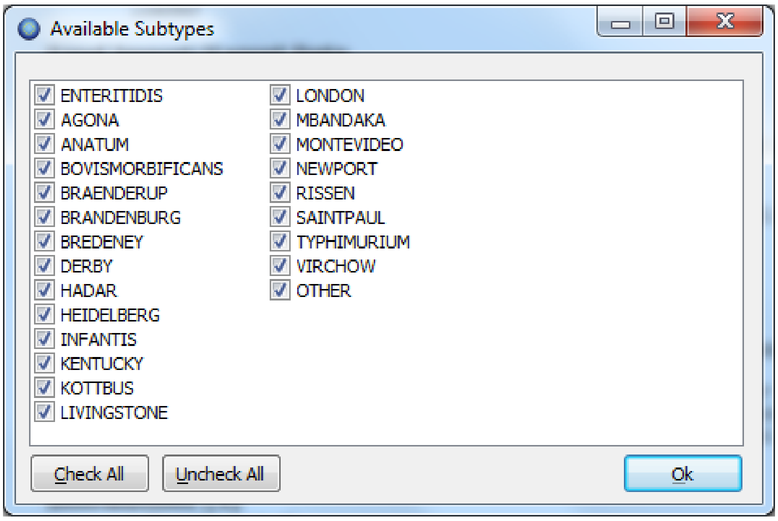
If new food sources have been added during data import, these food sources will be listed here.

By default, EFSA_SAM automatically selects all food sources present in the dataset, but it is possible to deselect food sources, if needed.

It should be noted that models require a minimum number of food sources to be able to run and provide reliable results. For the TT-SAM and Single MS model, the minimum number of food sources is three.

4.3. Selecting the bacteria subtypes

Here the bacteria subtypes are selected.



☑ ENTERITIDIS
☑ AGONA
☑ ANATUM
☑ BOVISMORBIFICANS
☑ BRAENDERUP
☑ BRANDENBURG
☑ BREDENEY
☑ DERBY
☑ HADAR
☑ HEIDELBERG
☑ INFANTIS
☑ KENTUCKY
☑ KOTTBUS
☑ LIVINGSTONE
☑ LONDON
☑ MBANDAKA
☑ MONTEVIDEO
☑ NEWPORT
☑ RISSEN
☑ SAINTPAUL
☑ TYPHIMURIUM
☑ VIRCHOW
☑ OTHER

Only subtypes present in the imported data (either in Food source prevalences or Reported human cases) will be shown.

By default, EFSA_SAM automatically selects all subtypes present in the dataset, but it is possible to deselect subtypes, if needed.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

It should be noted that models require a minimum number of subtypes to be able to run and provide reliable results (around 10 subtypes).

5. Model settings

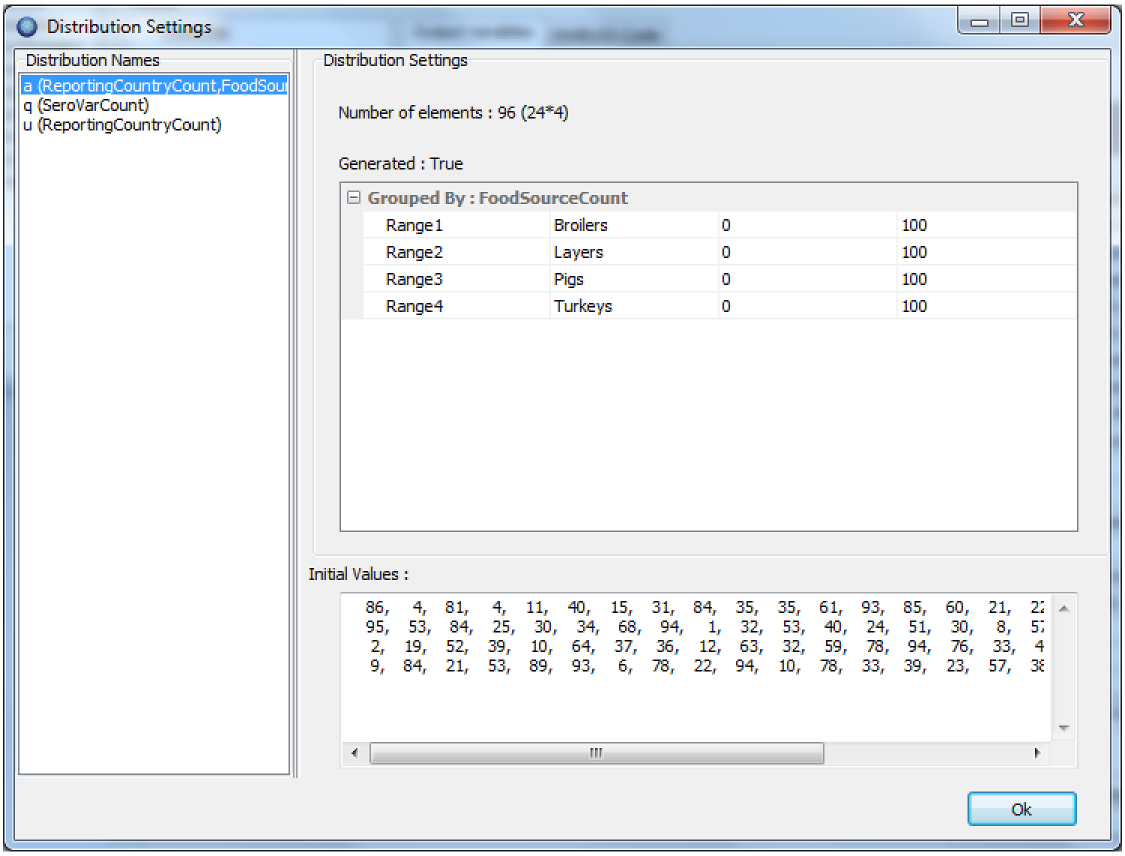
These are the settings for the model, where the user can check or revise the underreporting factors for human cases if relevant for the model selected, and add ”What If”-*scenarios*.

5.1. Distribution settings

The number of distributions presents, depends of the model selected. In the *Distribution Settings* window for the priors, the ranges for each distribution (grouped or ungrouped) can be set independently.

In the example below, you can see that distribution *a* is a two dimensional array over *Reporting Countries* and *Food Sources.*

You can also see that the dimension is grouped by *Food Source*. Since four food sources have been chosen earlier, four lines appear in the window – one for each food source.



Distribution Names

a (ReportingCountryCount,FoodSou
q (SeroVarCount)
u (ReportingCountryCount)

Distribution Settings

Number of elements : 96 (24*4)

Generated : True

Grouped By : FoodSourceCount

Range1	Broilers	0	100
Range2	Layers	0	100
Range3	Pigs	0	100
Range4	Turkeys	0	100

Initial Values :

```
86,  4, 81,  4, 11, 40, 15, 31, 84, 35, 35, 61, 93, 85, 60, 21, 2
95, 53, 84, 25, 30, 34, 68, 94,  1, 32, 53, 40, 24, 51, 30,  8, 5
 2, 19, 52, 39, 10, 64, 37, 36, 12, 63, 32, 59, 78, 94, 76, 33, 4
 9, 84, 21, 53, 89, 93,  6, 78, 22, 94, 10, 78, 33, 39, 23, 57, 3
```

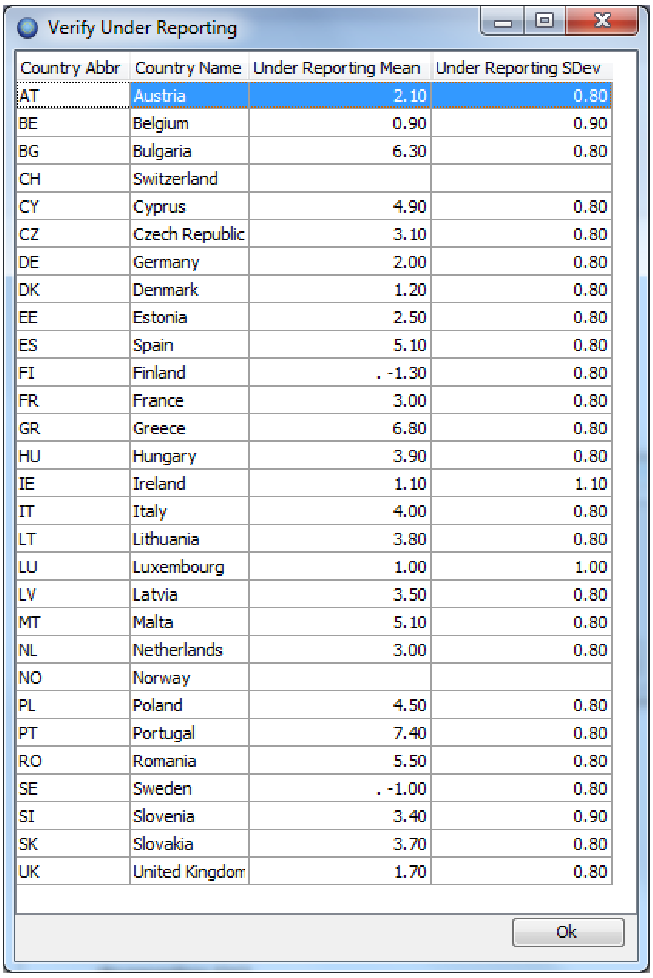
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

The *Dimension distributions* window will be available when the baseline analysis has been run, in order to adjust the ranges, if the *Density plots* show that the posterior distributions are out of range (how to check this is explained later).

5.2. Underreporting factors for human cases

If the model selected supports *underreporting factors*, this window will be available.

Here it is possible to verify and adjust the factors (distribution means and standard deviations) used. The underreporting factors available as default in the EFSA_SAM interface are based on the methodologies published by Havelaar et al. (2012)[8].



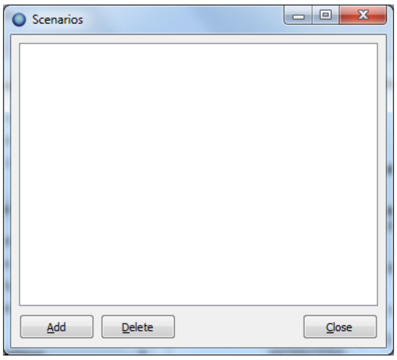
Country Abbr	Country Name	Under Reporting Mean	Under Reporting SDev
AT	Austria	2.10	0.80
BE	Belgium	0.90	0.90
BG	Bulgaria	6.30	0.80
CH	Switzerland		
CY	Cyprus	4.90	0.80
CZ	Czech Republic	3.10	0.80
DE	Germany	2.00	0.80
DK	Denmark	1.20	0.80
EE	Estonia	2.50	0.80
ES	Spain	5.10	0.80
FI	Finland	. -1.30	0.80
FR	France	3.00	0.80
GR	Greece	6.80	0.80
HU	Hungary	3.90	0.80
IE	Ireland	1.10	1.10
IT	Italy	4.00	0.80
LT	Lithuania	3.80	0.80
LU	Luxembourg	1.00	1.00
LV	Latvia	3.50	0.80
MT	Malta	5.10	0.80
NL	Netherlands	3.00	0.80
NO	Norway		
PL	Poland	4.50	0.80
PT	Portugal	7.40	0.80
RO	Romania	5.50	0.80
SE	Sweden	. -1.00	0.80
SI	Slovenia	3.40	0.90
SK	Slovakia	3.70	0.80
UK	United Kingdom	1.70	0.80

[8] Havelaar AH, Ivarsson S, Löfdahl M and Nauta MJ, in press. Estimating the true incidence of campylobacteriosis and salmonellosis in the EU, 2009. Epidemiology and Infection, published online 13 April 2012. DOI: http://dx.doi.org/10.1017/S0950268812000568

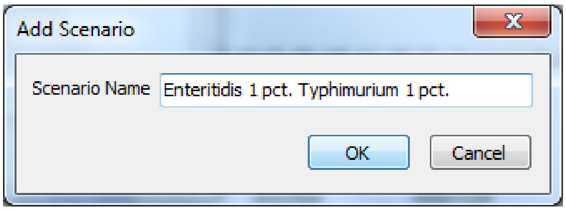
6. Scenarios

Scenarios can be added to EU multi member state models by simply copying the *baseline* data and settings, and allowing the user to change target prevalence data, and then execute different scenarios in WinBUGS.

It is possible to set target prevalences for either individual subtypes or for groups of subtypes. You can make any number of scenarios.



In order to create a new scenario, press add, and a dialog box will ask you to give the scenario a name.

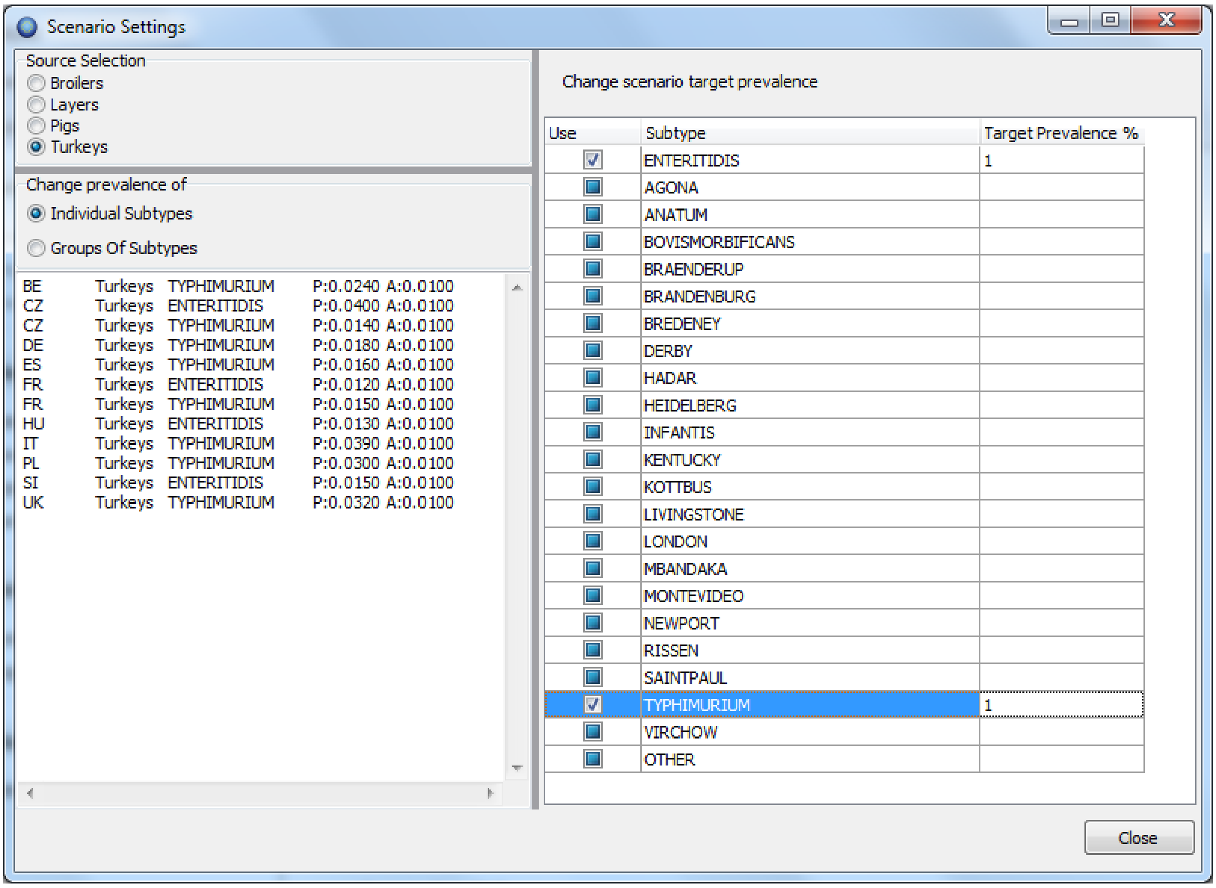


Scenario Name: Enteritidis 1 pct. Typhimurium 1 pct.

It is important to enter a meaningful name here, as this name will be used for identifying the scenarios afterwards. Press Ok.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

A new *Scenario Settings* window appears.



Scenario Settings

Source Selection
○ Broilers
○ Layers
○ Pigs
● Turkeys

Change prevalence of
● Individual Subtypes
○ Groups Of Subtypes

```
BE    Turkeys  TYPHIMURIUM    P:0.0240 A:0.0100
CZ    Turkeys  ENTERITIDIS    P:0.0400 A:0.0100
CZ    Turkeys  TYPHIMURIUM    P:0.0140 A:0.0100
DE    Turkeys  TYPHIMURIUM    P:0.0180 A:0.0100
ES    Turkeys  TYPHIMURIUM    P:0.0160 A:0.0100
FR    Turkeys  ENTERITIDIS    P:0.0120 A:0.0100
FR    Turkeys  TYPHIMURIUM    P:0.0150 A:0.0100
HU    Turkeys  ENTERITIDIS    P:0.0130 A:0.0100
IT    Turkeys  TYPHIMURIUM    P:0.0390 A:0.0100
PL    Turkeys  TYPHIMURIUM    P:0.0300 A:0.0100
SI    Turkeys  ENTERITIDIS    P:0.0150 A:0.0100
UK    Turkeys  TYPHIMURIUM    P:0.0320 A:0.0100
```

Change scenario target prevalence

Use	Subtype	Target Prevalence %
☑	ENTERITIDIS	1
□	AGONA	
□	ANATUM	
□	BOVISMORBIFICANS	
□	BRAENDERUP	
□	BRANDENBURG	
□	BREDENEY	
□	DERBY	
□	HADAR	
□	HEIDELBERG	
□	INFANTIS	
□	KENTUCKY	
□	KOTTBUS	
□	LIVINGSTONE	
□	LONDON	
□	MBANDAKA	
□	MONTEVIDEO	
□	NEWPORT	
□	RISSEN	
□	SAINTPAUL	
☑	TYPHIMURIUM	1
□	VIRCHOW	
□	OTHER	

Here you select the food source, and whether you want to change the target prevalence for individual subtypes or for groups of subtypes.

6.1. Individual subtypes

As shown in the screenshot above, you will see the subtypes present in the analysis in the right side table.

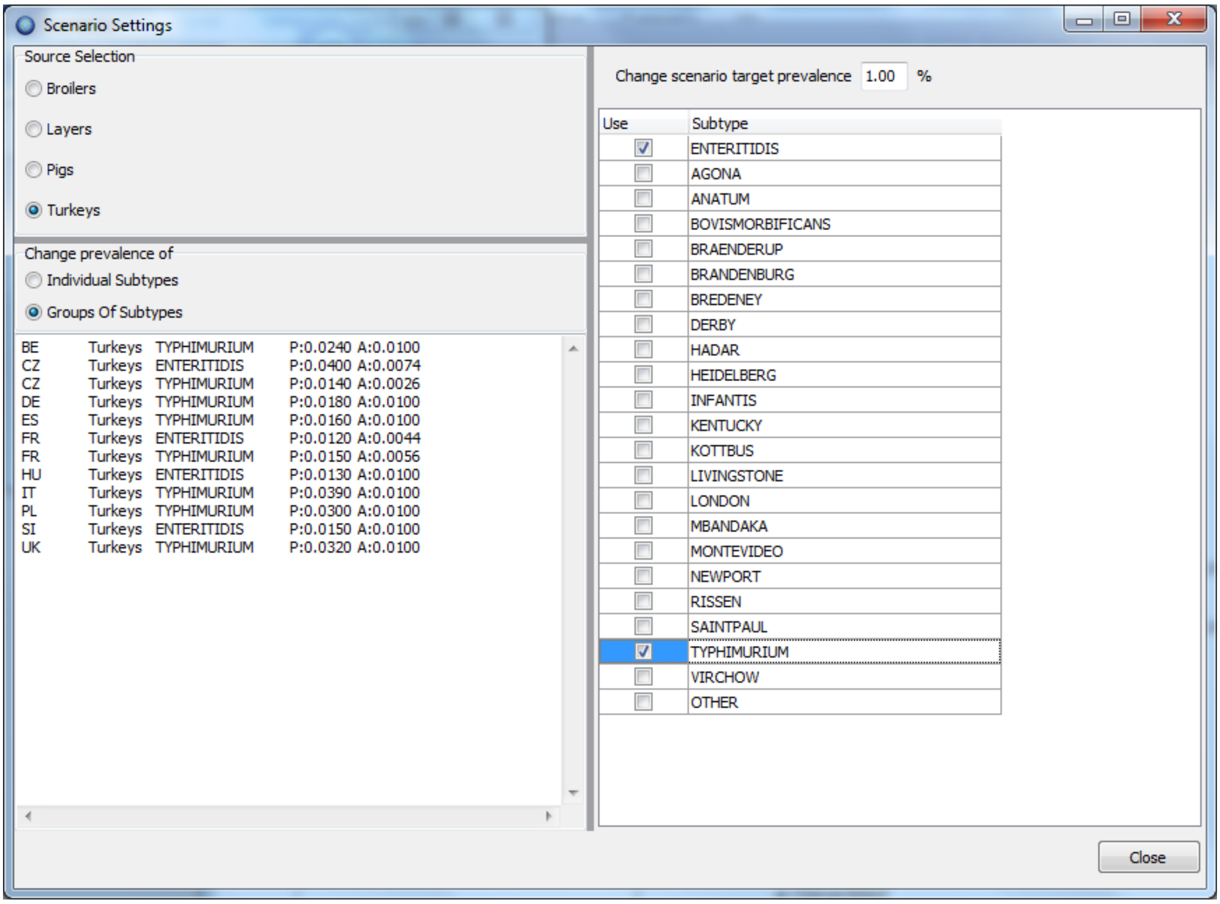
Click the check box and set a target prevalence value for that specific subtype. Multiple subtypes can be set as needed, each with a different target prevalence. It is, however, only possible to set target prevalence values for one food source at the time in each scenario.

In the window to the left, a list of affected subtypes will appear, telling the actual prevalence (P) and the adjusted prevalence (A).

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

6.2. Groups of subtypes

In the table to the right, you will see the subtypes present in the analysis. An editbox above the table enables you to enter the target prevalence for a group of subtype (the subtypes selected using the check boxes)



Source Selection

○ Broilers

○ Layers

○ Pigs

● Turkeys

Change prevalence of

○ Individual Subtypes

● Groups Of Subtypes

```
BE   Turkeys  TYPHIMURIUM   P:0.0240 A:0.0100
CZ   Turkeys  ENTERITIDIS   P:0.0400 A:0.0074
CZ   Turkeys  TYPHIMURIUM   P:0.0140 A:0.0026
DE   Turkeys  TYPHIMURIUM   P:0.0180 A:0.0100
ES   Turkeys  TYPHIMURIUM   P:0.0160 A:0.0100
FR   Turkeys  ENTERITIDIS   P:0.0120 A:0.0044
FR   Turkeys  TYPHIMURIUM   P:0.0150 A:0.0056
HU   Turkeys  ENTERITIDIS   P:0.0130 A:0.0100
IT   Turkeys  TYPHIMURIUM   P:0.0390 A:0.0100
PL   Turkeys  TYPHIMURIUM   P:0.0300 A:0.0100
SI   Turkeys  ENTERITIDIS   P:0.0150 A:0.0100
UK   Turkeys  TYPHIMURIUM   P:0.0320 A:0.0100
```

Change scenario target prevalence 1.00 %

Use	Subtype
☑	ENTERITIDIS
☐	AGONA
☐	ANATUM
☐	BOVISMORBIFICANS
☐	BRAENDERUP
☐	BRANDENBURG
☐	BREDENEY
☐	DERBY
☐	HADAR
☐	HEIDELBERG
☐	INFANTIS
☐	KENTUCKY
☐	KOTTBUS
☐	LIVINGSTONE
☐	LONDON
☐	MBANDAKA
☐	MONTEVIDEO
☐	NEWPORT
☐	RISSEN
☐	SAINTPAUL
☑	TYPHIMURIUM
☐	VIRCHOW
☐	OTHER

In the window to the left, a list of affected subtypes will appear, telling the actual prevalence (P) and the adjusted prevalence (A).

Due to minor rounding errors the calculated new grouped prevalence may not for some countries equal exactly the selected target value (+/- 0.0001).

7. Running the analysis

When all settings and choices have been made, it is time to run the analysis for the baseline.

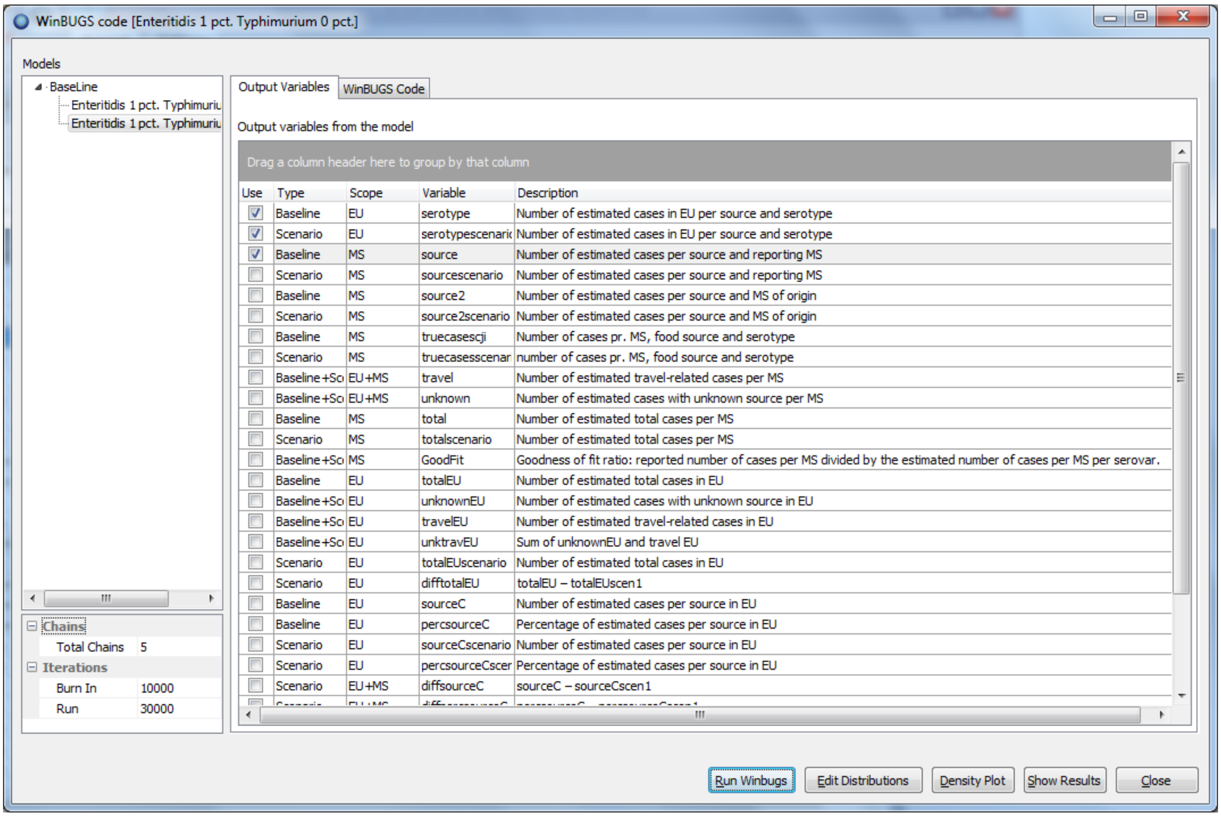
In the tree view to the left, you can see the baseline, and if you have chosen to create scenarios, these will also be present in the view.

In order to run the baseline, you must select the baseline, likewise, if you want to run a scenario, select that scenario, and set the variables.

In order to avoid using and maintaining too many different models (WinBUGS source code), the EFSA_SAM is built in a way, that it will run both a baseline and a scenario for each run. All scenarios use the same WinBUGS code.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

When selecting the baseline, the baseline data will be used and the results will reflect the results of the baseline. When selecting a scenario, both the baseline data and the scenario data (with the adjusted prevalences) will be used and results from both the baseline and the selected scenario will be outputted.


WinBUGS code [Enteritidis 1 pct. Typhimurium 0 pct.]
Models
BaseLine
Enteritidis 1 pct. Typhimuriu
Enteritidis 1 pct. Typhimuriu
Output Variables
WinBUGS Code
Drag a column header here to group by that column
Run Winbugs
Edit Distributions
Density Plot
Show Results
Close

Output variables from the model

Use	Type	Scope	Variable	Description
☑	Baseline	EU	serotype	Number of estimated cases in EU per source and serotype
☑	Scenario	EU	serotypescenari	Number of estimated cases in EU per source and serotype
☑	Baseline	MS	source	Number of estimated cases per source and reporting MS
□	Scenario	MS	sourcescenario	Number of estimated cases per source and reporting MS
□	Baseline	MS	source2	Number of estimated cases per source and MS of origin
□	Scenario	MS	source2scenario	Number of estimated cases per source and MS of origin
□	Baseline	MS	truecasescji	Number of cases pr. MS, food source and serotype
□	Scenario	MS	truecasesscenar	number of cases pr. MS, food source and serotype
□	Baseline+Sc	EU+MS	travel	Number of estimated travel-related cases per MS
□	Baseline+Sc	EU+MS	unknown	Number of estimated cases with unknown source per MS
□	Baseline	MS	total	Number of estimated total cases per MS
□	Scenario	MS	totalscenario	Number of estimated total cases per MS
□	Baseline+Sc	MS	GoodFit	Goodness of fit ratio: reported number of cases per MS divided by the estimated number of cases per MS per serovar.
□	Baseline	EU	totalEU	Number of estimated total cases in EU
□	Baseline+Sc	EU	unknownEU	Number of estimated cases with unknown source in EU
□	Baseline+Sc	EU	travelEU	Number of estimated travel-related cases in EU
□	Baseline+Sc	EU	unktravEU	Sum of unknownEU and travel EU
□	Scenario	EU	totalEUscenario	Number of estimated total cases in EU
□	Scenario	EU	difftotalEU	totalEU – totalEUscen1
□	Baseline	EU	sourceC	Number of estimated cases per source in EU
□	Baseline	EU	percsourceC	Percentage of estimated cases per source in EU
□	Scenario	EU	sourceCscenario	Number of estimated cases per source in EU
□	Scenario	EU	percsourceCscer	Percentage of estimated cases per source in EU
□	Scenario	EU+MS	diffsourceC	sourceC – sourceCscen1

Chains	
Total Chains	5
Iterations	
Burn In	10000
Run	30000

Now you can select the output variable that you want to obtain results from. It is always a good idea to select the “GoodFit” variable, as this will give an indication of how well the model is performing i.e. able to fit the data (see also the section on model convergence and goodness of fit below).

Here, you also enter the values for the number of Markov chains, the number of Burn In iterations, and the number of Run iterations, that WinBUGS will use during execution.

Note: When running WinBUGS in script mode, as we do here, the total number of iterations run is the Burn In iterations PLUS the Run Iterations. When running in GUI-mode (using the WinBUGS user interface), the total number of iterations is the run-iterations, where the user can selecct how many of these should be counted as Burn In.

As a start, chains and iteration values will be set to a default value (chosen by the creator of the model) the first time you enter the baseline and scenarios. The number of iterations needed depends on the complexity of the model, but a Burn In of 10,000 and a number of iteration of 30,000 run in 3-5 chains appears appropriate for the EU *Salmonella* model (the demo data).

Since you might want to run the scenarios under different conditions, and wanting different data outputs, each scenario has its own set of variables and WinBUGS run-settings (Chains and Iterations).

Above the table with the variables, you can see two tabs. The Output Variables, and the WinBUGS code tab.

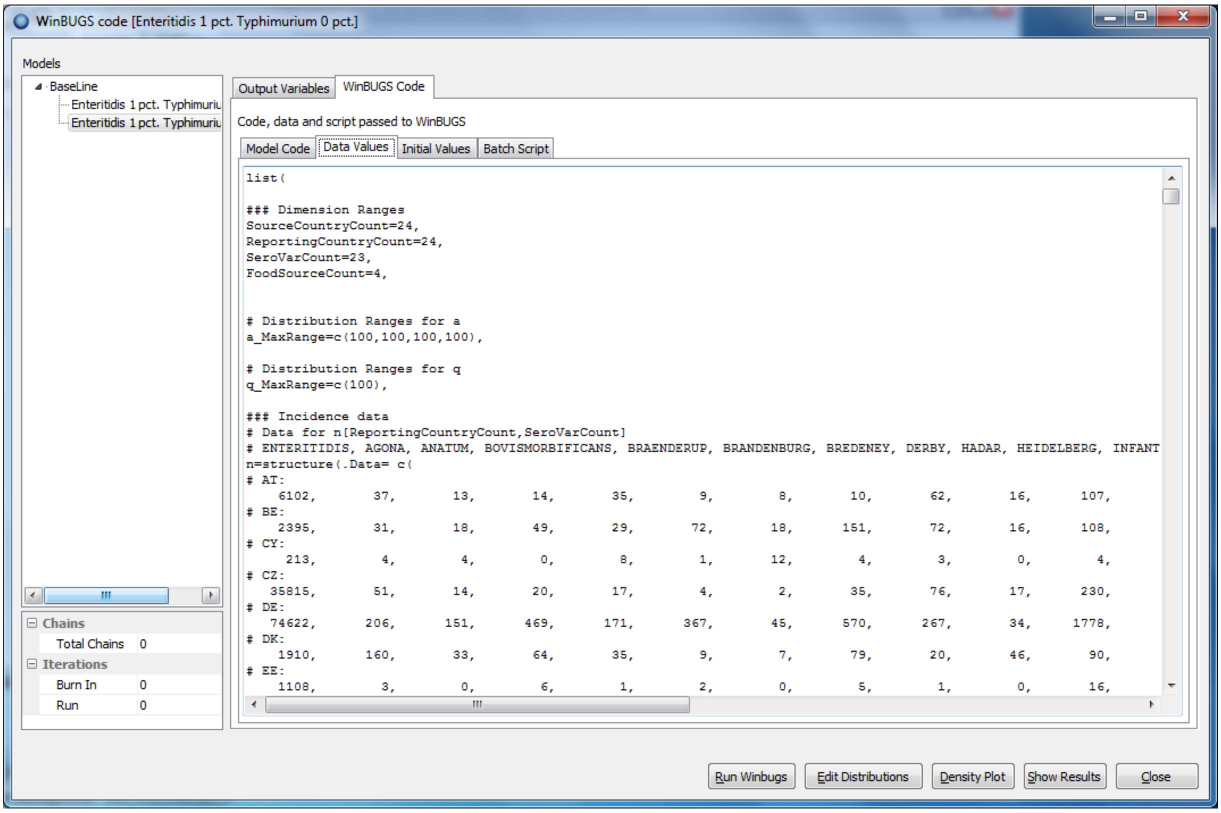
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

If you need to review the data, the code or script the EFSA_SAM has created, click on the WinBUGS code tab.

As shown in the following window, it is possible for advanced users to review and edit the model code. The code will not be shown unless the user chooses to show it.

If the model code is changed, these changes will be saved in the final analysis files, but it will not be saved into the original model settings in the database.

If the user wants to save the changes for future use, it will be possible to add a changed model to the model repository.


WinBUGS code [Enteritidis 1 pct. Typhimurium 0 pct.]
Models
BaseLine
Enteritidis 1 pct. Typhimuriu
Enteritidis 1 pct. Typhimuriu
Output Variables
WinBUGS Code
Code, data and script passed to WinBUGS
Model Code
Data Values
Initial Values
Batch Script
Run Winbugs
Edit Distributions
Density Plot
Show Results
Close

Chains
Total Chains 0
Iterations
Burn In 0
Run 0

```
list(

### Dimension Ranges
SourceCountryCount=24,
ReportingCountryCount=24,
SeroVarCount=23,
FoodSourceCount=4,


# Distribution Ranges for a
a_MaxRange=c(100,100,100,100),

# Distribution Ranges for q
q_MaxRange=c(100),

### Incidence data
# Data for n[ReportingCountryCount,SeroVarCount]
# ENTERITIDIS, AGONA, ANATUM, BOVISMORBIFICANS, BRAENDERUP, BRANDENBURG, BREDENEY, DERBY, HADAR, HEIDELBERG, INFANT
n=structure(.Data= c(
# AT:
    6102,       37,       13,       14,       35,        9,        8,       10,       62,       16,      107,
# BE:
    2395,       31,       18,       49,       29,       72,       18,      151,       72,       16,      108,
# CY:
     213,        4,        4,        0,        8,        1,       12,        4,        3,        0,        4,
# CZ:
   35815,       51,       14,       20,       17,        4,        2,       35,       76,       17,      230,
# DE:
   74622,      206,      151,      469,      171,      367,       45,      570,      267,       34,     1778,
# DK:
    1910,      160,       33,       64,       35,        9,        7,       79,       20,       46,       90,
# EE:
    1108,        3,        0,        6,        1,        2,        0,        5,        1,        0,       16,
```

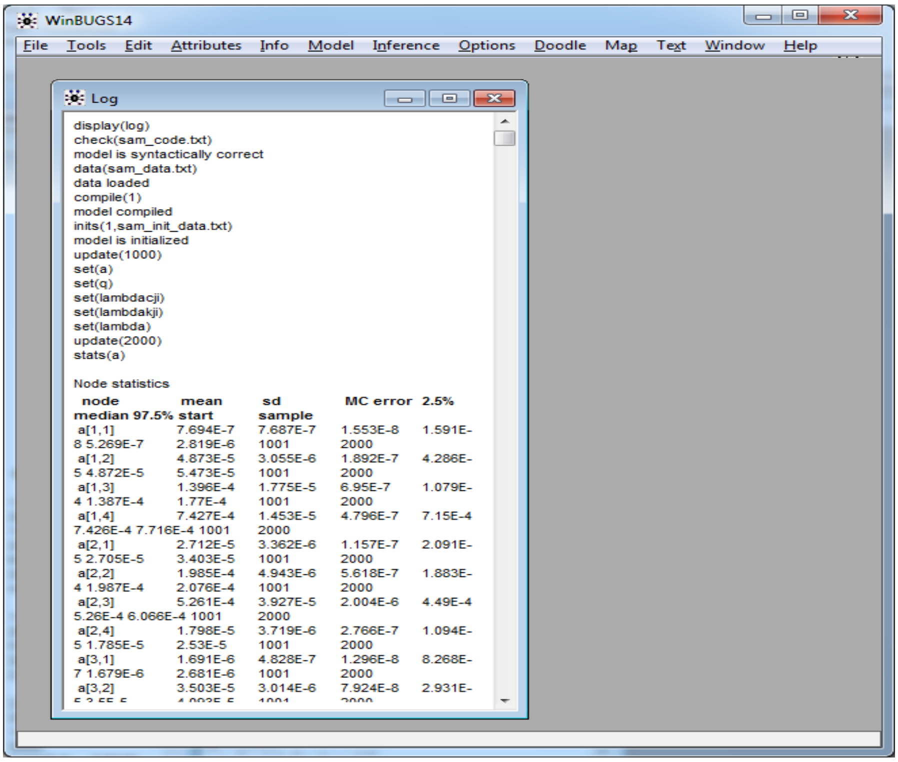
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

7.1. Model execution

Basically, all you have to do is to select the *Baseline,* select the output variables, set the number of chains and iterations and then press the *Run WinBUGS*-button.

Once these values have been set, the initial values for the distributions will be created and added to the file that WinBUGS uses for initial values. Once the initial values have been created, they are saved in the analysis file, and will not be changed, unless the dimensions or the distribution range are changed.

Then EFSA_SAM will generate all the files that WinBUGS needs in order to run the analysis, and start WinBUGS execution.



```
display(log)
check(sam_code.txt)
model is syntactically correct
data(sam_data.txt)
data loaded
compile(1)
model compiled
inits(1,sam_init_data.txt)
model is initialized
update(1000)
set(a)
set(q)
set(lambdacji)
set(lambdakji)
set(lambda)
update(2000)
stats(a)

Node statistics
 node            mean        sd           MC error  2.5%
median 97.5% start          sample
 a[1,1]          7.694E-7    7.687E-7     1.553E-8    1.591E-
8 5.269E-7       2.819E-6    1001         2000
 a[1,2]          4.873E-5    3.055E-6     1.892E-7    4.286E-
5 4.872E-5       5.473E-5    1001         2000
 a[1,3]          1.396E-4    1.775E-5     6.95E-7     1.079E-
4 1.387E-4       1.77E-4     1001         2000
 a[1,4]          7.427E-4    1.453E-5     4.796E-7    7.15E-4
7.426E-4 7.716E-4 1001       2000
 a[2,1]          2.712E-5    3.362E-6     1.157E-7    2.091E-
5 2.705E-5       3.403E-5    1001         2000
 a[2,2]          1.985E-4    4.943E-6     5.618E-7    1.883E-
4 1.987E-4       2.076E-4    1001         2000
 a[2,3]          5.261E-4    3.927E-5     2.004E-6    4.49E-4
5.26E-4 6.066E-4 1001        2000
 a[2,4]          1.798E-5    3.719E-6     2.766E-7    1.094E-
5 1.785E-5       2.53E-5     1001         2000
 a[3,1]          1.691E-6    4.828E-7     1.296E-8    8.268E-
7 1.679E-6       2.681E-6    1001         2000
 a[3,2]          3.503E-5    3.014E-6     7.924E-8    2.931E-
5 3.55E-5        4.093E-5    1001         2000
```

When WinBUGS has finished running the analysis, it will automatically close and return to EFSA_SAM. The run time depends on the complexity of the model, the number of iterations and chains, and the power of the PC running the analysis, but it can easily take several hours using the recommended values and settings above.

If an error occurs during execution, EFSA_SAM will detect this and ask if the user wants to view the WinBUGS log, which can provide more information about the problem. For more serious errors, WinBUGS will stop execution with a so-called TRAP-window. This will happen rarely, and can be caused by e.g. values of zero in distributions or errors in the WinBUGS code written by the creator of the model.

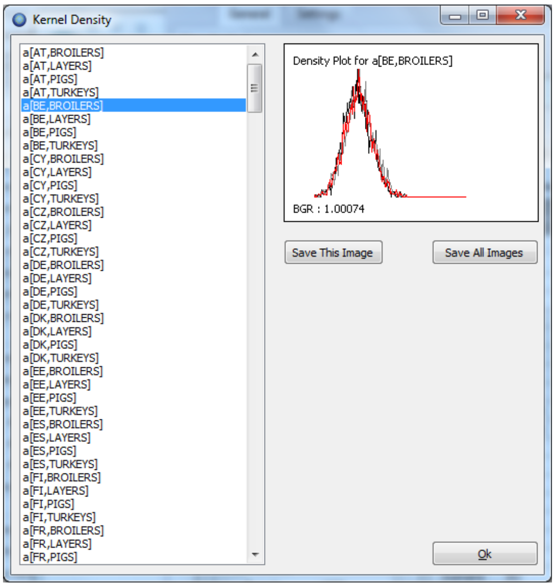
If the rare occasion arises, press the *"Recalculate Initial Values"*-button.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

8. Monitoring of convergence and goodness of fit

If no problems occurred during the execution, you can now press *Kernel Density*, and a window displaying the *Kernel density plots* for the food-source-dependent factors (a) and subtype-dependent factors (q) will be shown.

All the plot names are listed to the left, and by clicking the name, the plots for that particular variable name is displayed in the right side. There will be one curve for each of the chains run in the model.



a[AT,BROILERS]
a[AT,LAYERS]
a[AT,PIGS]
a[AT,TURKEYS]
a[BE,BROILERS]
a[BE,LAYERS]
a[BE,PIGS]
a[BE,TURKEYS]
a[CY,BROILERS]
a[CY,LAYERS]
a[CY,PIGS]
a[CY,TURKEYS]
a[CZ,BROILERS]
a[CZ,LAYERS]
a[CZ,PIGS]
a[CZ,TURKEYS]
a[DE,BROILERS]
a[DE,LAYERS]
a[DE,PIGS]
a[DE,TURKEYS]
a[DK,BROILERS]
a[DK,LAYERS]
a[DK,PIGS]
a[DK,TURKEYS]
a[EE,BROILERS]
a[EE,LAYERS]
a[EE,PIGS]
a[EE,TURKEYS]
a[ES,BROILERS]
a[ES,LAYERS]
a[ES,PIGS]
a[ES,TURKEYS]
a[FI,BROILERS]
a[FI,LAYERS]
a[FI,PIGS]
a[FI,TURKEYS]
a[FR,BROILERS]
a[FR,LAYERS]
a[FR,PIGS]

If needed, the plots can be saved to bitmap files (bmp). You can save either the plot shown, or choose to save all plots. When saving plots you are asked for a location folder for the plots. Each file will automatically be given the name of the variable plotted.

A first step to check how the model is performing is to examine these Kernel density plots. They should depict distributions that at the right side end in a smooth curve approximating zero. If the density distribution is cut off to the right the ranges of the prior distribution appear inappropriate and need to be extended (see the technical report for more information on this). EFSA_SAM will display a small warning (three stars) if the density peak is off centered. A warning will be displayed if the peak is in the right 15 percent of the range. Even though warnings are displayed, it is a good idea to verify that all plots are ok.

If changes are made to the distribution ranges, new initial values for the ranges changed will be made, and the *baseline*-analysis must be run again.

The use of more Markov chains and monitoring that these chains after a specified number of iterations are ending up at the same results, is another appropriate way of monitoring for convergence.

A very simple way of monitoring convergence is by visually looking at the Kernel density plot for the food-source-dependent factors (a) and subtype-dependent factors (q) for each chain. If these appear to be overlying, the model seems to have converged well.

A more formal approach is to calculate the bgr-diagnostics. Here convergence is considered to have occurred when the variance across all chains (B) is no larger than the variance within each individual chain (W), and when the chains had reached a stable level. EFSA_SAM makes a calculation of the

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

bgr-diagnostic based on the results of *a* and *q* for the second half of the iterations resulting in a single value for R for each posterior distribution. Convergence may be assumed for practical purposes if R < 1.05 (WinBUGS user manual).

Finally, the predictive ability – or goodness of fit - of the model can be assessed by estimating the ratio between the observed human cases (sporadic human cases reported in each country) and the number of cases predicted by the model. A ratio around 1 indicates a good fit. A plot of the GoodFit values from each country can indicate if the model for some countries is predicting poorly as compared to the observed data. Poor fit is often related to data quality and the user may, therefore, want to exclude specific countries in order to evaluate the influence of these countries on the overall model results.

For the single MS model, the goodness of fit can be assessed by comparing the variables lambdaexp (expected number of cases per subtype and year) with n (the reported number of cases per subtype and year).

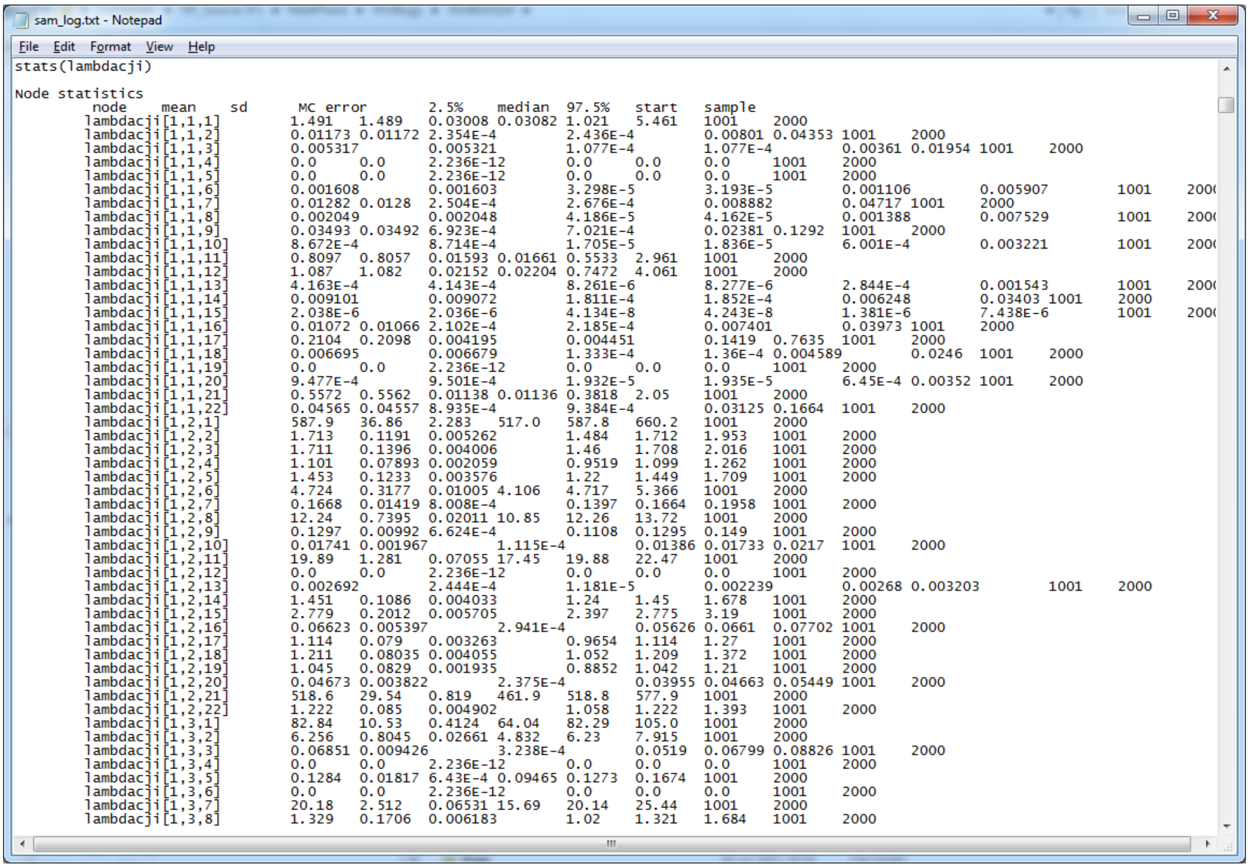
When the user is satisfied with the performance of the model, it is time to run the *Scenarios*.

This time, select a scenario in the tree view, set the output variables and WinBUGS run-settings and press *Run WinBUGS* just like you did when you ran the *Baseline.*

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

9. The output

After the analysis runs, the data produced by WinBUGS will be imported to EFSA_SAM.

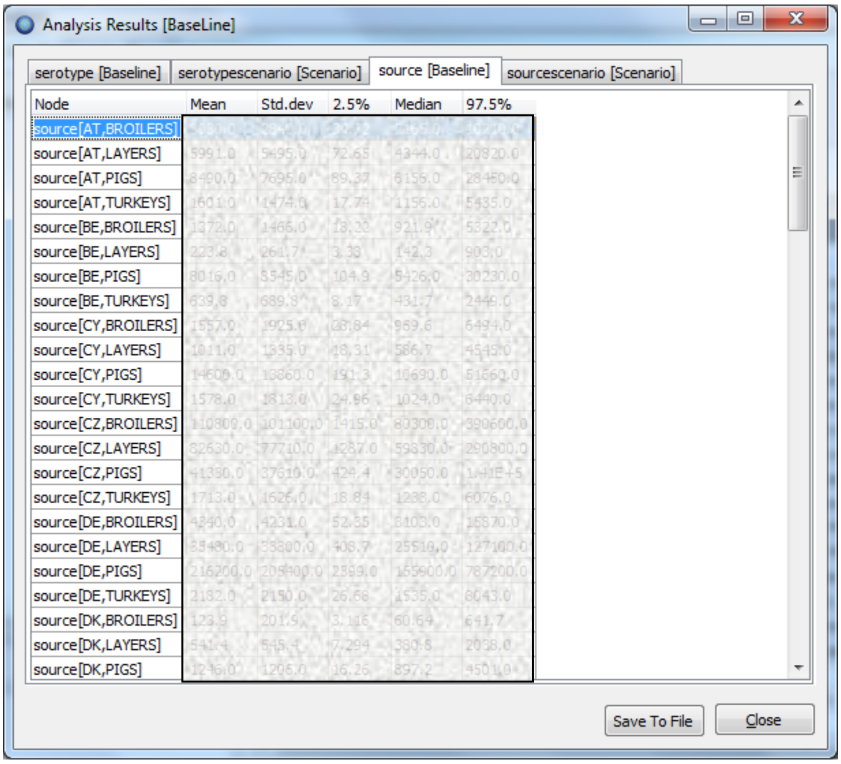


```
stats(lambdacji)

Node statistics
         node    mean    sd      MC error        2.5%    median  97.5%   start   sample
        lambdacji[1,1,1]        1.491   1.489   0.03008 0.03082 1.021   5.461   1001    2000
        lambdacji[1,1,2]        0.01173 0.01172 2.354E-4        2.436E-4        0.00801 0.04353 1001    2000
        lambdacji[1,1,3]        0.005317        0.005321        1.077E-4        1.077E-4        0.00361 0.01954 1001    2000
        lambdacji[1,1,4]        0.0     0.0     2.236E-12       0.0     0.0     0.0     1001    2000
        lambdacji[1,1,5]        0.0     0.0     2.236E-12       0.0     0.0     0.0     1001    2000
        lambdacji[1,1,6]        0.001608        0.001603        3.298E-5        3.193E-5        0.001106        0.005907        1001    200
        lambdacji[1,1,7]        0.01282 0.0128  2.504E-4        2.676E-4        0.008882        0.04717 1001    2000
        lambdacji[1,1,8]        0.002049        0.002048        4.186E-5        4.162E-5        0.001388        0.007529        1001    200
        lambdacji[1,1,9]        0.03493 0.03492 6.923E-4        7.021E-4        0.02381 0.1292  1001    2000
        lambdacji[1,1,10]       8.672E-4        8.714E-4        1.705E-5        1.836E-5        6.001E-4        0.003221        1001    200
        lambdacji[1,1,11]       0.8097  0.8057  0.01593 0.01661 0.5533  2.961   1001    2000
        lambdacji[1,1,12]       1.087   1.082   0.02152 0.02204 0.7472  4.061   1001    2000
        lambdacji[1,1,13]       4.163E-4        4.143E-4        8.261E-6        8.277E-6        2.844E-4        0.001543        1001    200
        lambdacji[1,1,14]       0.009101        0.009072        1.811E-4        1.852E-4        0.006248        0.03403 1001    2000
        lambdacji[1,1,15]       2.038E-6        2.036E-6        4.134E-8        4.243E-8        1.381E-6        7.438E-6        1001    200
        lambdacji[1,1,16]       0.01072 0.01066 2.102E-4        2.185E-4        0.007401        0.03973 1001    2000
        lambdacji[1,1,17]       0.2104  0.2098  0.004195        0.004451        0.1419  0.7635  1001    2000
        lambdacji[1,1,18]       0.006695        0.006679        1.333E-4        1.36E-4 0.004589        0.0246  1001    2000
        lambdacji[1,1,19]       0.0     0.0     2.236E-12       0.0     0.0     0.0     1001    2000
        lambdacji[1,1,20]       9.477E-4        9.501E-4        1.932E-5        1.935E-5        6.45E-4 0.00352 1001    2000
        lambdacji[1,1,21]       0.5572  0.5562  0.01138 0.01136 0.3818  2.05    1001    2000
        lambdacji[1,1,22]       0.04565 0.04557 8.935E-4        9.384E-4        0.03125 0.1664  1001    2000
        lambdacji[1,2,1]        587.9   36.86   2.283   517.0   587.8   660.2   1001    2000
        lambdacji[1,2,2]        1.713   0.1191  0.005262        1.484   1.712   1.953   1001    2000
        lambdacji[1,2,3]        1.711   0.1396  0.004006        1.46    1.708   2.016   1001    2000
        lambdacji[1,2,4]        1.101   0.07893 0.002059        0.9519  1.099   1.262   1001    2000
        lambdacji[1,2,5]        1.453   0.1233  0.003576        1.22    1.449   1.709   1001    2000
        lambdacji[1,2,6]        4.724   0.3177  0.01005 4.106   4.717   5.366   1001    2000
        lambdacji[1,2,7]        0.1668  0.01419 8.008E-4        0.1397  0.1664  0.1958  1001    2000
        lambdacji[1,2,8]        12.24   0.7395  0.02011 10.85   12.26   13.72   1001    2000
        lambdacji[1,2,9]        0.1297  0.00992 6.624E-4        0.1108  0.1295  0.149   1001    2000
        lambdacji[1,2,10]       0.01741 0.001967        1.115E-4        0.01386 0.01733 0.0217  1001    2000
        lambdacji[1,2,11]       19.89   1.281   0.07055 17.45   19.88   22.47   1001    2000
        lambdacji[1,2,12]       0.0     0.0     2.236E-12       0.0     0.0     0.0     1001    2000
        lambdacji[1,2,13]       0.002692        2.444E-4        1.181E-5        0.002239        0.00268 0.003203        1001    2000
        lambdacji[1,2,14]       1.451   0.1086  0.004033        1.24    1.45    1.678   1001    2000
        lambdacji[1,2,15]       2.779   0.2012  0.005705        2.397   2.775   3.19    1001    2000
        lambdacji[1,2,16]       0.06623 0.005397        2.941E-4        0.05626 0.0661  0.07702 1001    2000
        lambdacji[1,2,17]       1.114   0.079   0.003263        0.9654  1.114   1.27    1001    2000
        lambdacji[1,2,18]       1.211   0.08035 0.004055        1.052   1.209   1.372   1001    2000
        lambdacji[1,2,19]       1.045   0.0829  0.001935        0.8852  1.042   1.21    1001    2000
        lambdacji[1,2,20]       0.04673 0.003822        2.375E-4        0.03955 0.04663 0.05449 1001    2000
        lambdacji[1,2,21]       518.6   29.54   0.819   461.9   518.8   577.9   1001    2000
        lambdacji[1,2,22]       1.222   0.085   0.004902        1.058   1.222   1.393   1001    2000
        lambdacji[1,3,1]        82.84   10.53   0.4124  64.04   82.29   105.0   1001    2000
        lambdacji[1,3,2]        6.256   0.8045  0.02661 4.832   6.23    7.915   1001    2000
        lambdacji[1,3,3]        0.06851 0.009426        3.238E-4        0.0519  0.06799 0.08826 1001    2000
        lambdacji[1,3,4]        0.0     0.0     2.236E-12       0.0     0.0     0.0     1001    2000
        lambdacji[1,3,5]        0.1284  0.01817 6.43E-4 0.09465 0.1273  0.1674  1001    2000
        lambdacji[1,3,6]        0.0     0.0     2.236E-12       0.0     0.0     0.0     1001    2000
        lambdacji[1,3,7]        20.18   2.512   0.06531 15.69   20.14   25.44   1001    2000
        lambdacji[1,3,8]        1.329   0.1706  0.006183        1.02    1.321   1.684   1001    2000
```

As you can see in the window above, the format of the output that WinBUGS makes, is not very easy to read or interpret. Therefore it will be re-arranged into a user friendly tabular format.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

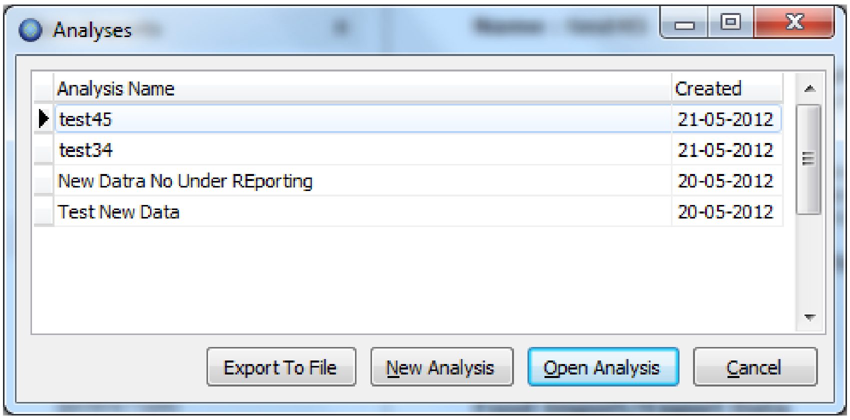

Analysis Results [BaseLine]
serotype [Baseline]
serotypescenario [Scenario]
source [Baseline]
sourcescenario [Scenario]
Save To File
Close

Node	Mean	Std.dev	2.5%	Median	97.5%
source[AT,BROILERS]	3030.0	2840.0	34.72	2165.0	10210.0
source[AT,LAYERS]	5991.0	5495.0	72.65	4344.0	20820.0
source[AT,PIGS]	8490.0	7695.0	89.37	6156.0	28450.0
source[AT,TURKEYS]	1601.0	1474.0	17.74	1156.0	5435.0
source[BE,BROILERS]	1372.0	1466.0	18.22	921.9	5322.0
source[BE,LAYERS]	223.8	261.7	3.38	142.3	903.0
source[BE,PIGS]	8016.0	8545.0	104.9	5426.0	30230.0
source[BE,TURKEYS]	639.8	689.8	8.17	431.7	2449.0
source[CY,BROILERS]	1557.0	1925.0	28.84	969.6	6494.0
source[CY,LAYERS]	1011.0	1335.0	18.31	586.7	4545.0
source[CY,PIGS]	14600.0	13860.0	191.3	10690.0	51660.0
source[CY,TURKEYS]	1578.0	1813.0	24.96	1024.0	6440.0
source[CZ,BROILERS]	110800.0	101100.0	1415.0	80300.0	390600.0
source[CZ,LAYERS]	82630.0	77710.0	1287.0	59830.0	290800.0
source[CZ,PIGS]	41330.0	37610.0	424.4	30050.0	1.41E+5
source[CZ,TURKEYS]	1713.0	1626.0	18.84	1238.0	6076.0
source[DE,BROILERS]	4340.0	4231.0	52.35	3103.0	15870.0
source[DE,LAYERS]	35480.0	33800.0	408.7	25510.0	127100.0
source[DE,PIGS]	216200.0	205400.0	2399.0	155900.0	787200.0
source[DE,TURKEYS]	2182.0	2150.0	26.68	1535.0	8043.0
source[DK,BROILERS]	123.9	201.9	3.116	60.64	641.7
source[DK,LAYERS]	541.4	545.4	7.294	380.8	2038.0
source[DK,PIGS]	1246.0	1206.0	16.26	897.2	4501.0

The data from the variables selected before running the model, is added to individual grids and tables. The node names are reformatted to contain the dimension values for easy identification of the data points. As seen in the screen shot above source[1,1] (the node name actually outputted from WinBUGS) is renamed to source[AT,Broilers], which gives the number of estimated human cases attributed to broilers in Austria.

All data including imported data, analysis settings, resulting data and model code will be saved in the database for later retrieval. This ensures that it is possible to see exactly how the results were produced years after it was done.

Also, it is possible to export all this as a zip-file. Each individual data file is also available. The resulting data values can also be exported as individual files for further analysis, in an easy to use format. To export an analysis as a zip-file, select *"Open Analysis"* and press *"Export to file"*. Then select a destination folder for your file and press *"Save"*



Analysis Name	Created
test45	21-05-2012
test34	21-05-2012
New Datra No Under REporting	20-05-2012
Test New Data	20-05-2012

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

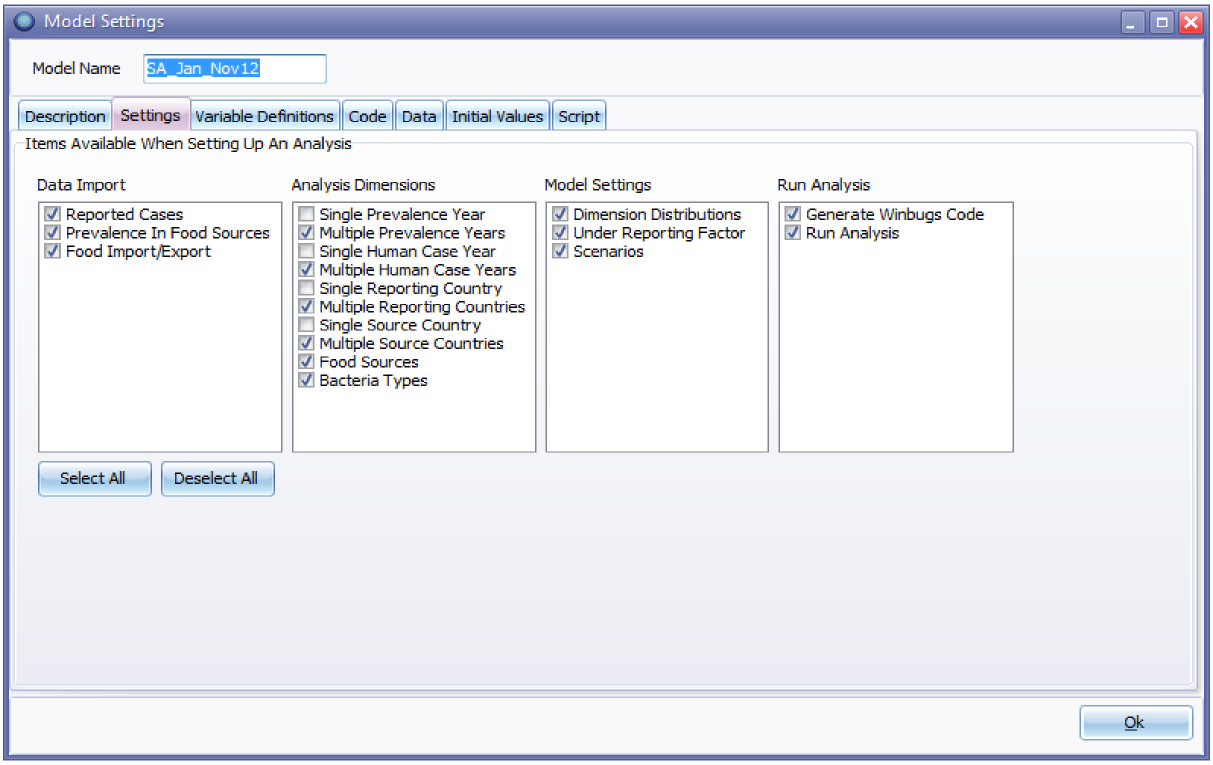
10. The advanced part of EFSA_SAM

This part of the manual is aimed for advanced users, needing to adapt EFSA_SAM to WinBUGS models.

Adding new models into EFSA_SAM is done in the *Model Settings* windows found in the *Settings Menu.*

10.1. Model settings

On the *settings* tab, it is possible to select which items are to be shown in the *sidebar*. Items not needed in order to run the model should be unchecked, making the user interface simpler and easier to use when actually running the analyses based on the model.



Model Name SA_Jan_Nov12

Items Available When Setting Up An Analysis

Data Import
- ☑ Reported Cases
- ☑ Prevalence In Food Sources
- ☑ Food Import/Export

Analysis Dimensions
- ☐ Single Prevalence Year
- ☑ Multiple Prevalence Years
- ☐ Single Human Case Year
- ☑ Multiple Human Case Years
- ☐ Single Reporting Country
- ☑ Multiple Reporting Countries
- ☐ Single Source Country
- ☑ Multiple Source Countries
- ☑ Food Sources
- ☑ Bacteria Types

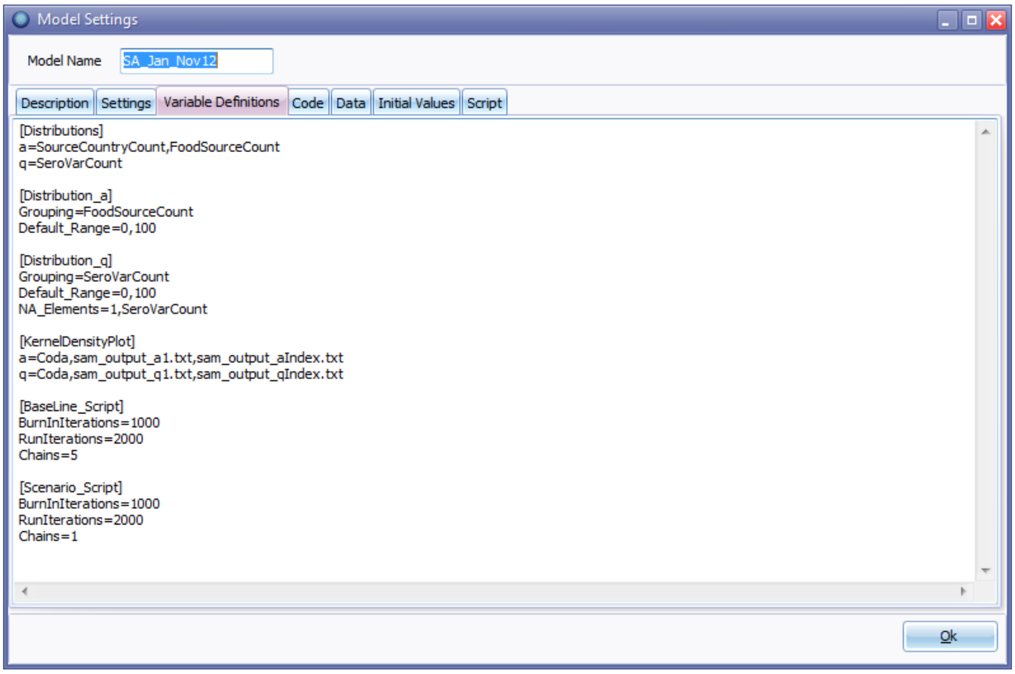
Model Settings
- ☑ Dimension Distributions
- ☑ Under Reporting Factor
- ☑ Scenarios

Run Analysis
- ☑ Generate Winbugs Code
- ☑ Run Analysis

10.2. Variable definitions

The variable definitions tab, is where the interface between EFSA_SAM and WinBUGS is defined.

These settings define *Distributions* their grouping, NA items and their default range values. This also includes which variables that are going to have *Kernel density plots*, default values for the *Baseline* and the *Scenarios* scripts (iterations and chains) etc.



```
[Distributions]
a=SourceCountryCount,FoodSourceCount
q=SeroVarCount

[Distribution_a]
Grouping=FoodSourceCount
Default_Range=0,100

[Distribution_q]
Grouping=SeroVarCount
Default_Range=0,100
NA_Elements=1,SeroVarCount

[KernelDensityPlot]
a=Coda,sam_output_a1.txt,sam_output_aIndex.txt
q=Coda,sam_output_q1.txt,sam_output_qIndex.txt

[BaseLine_Script]
BurnInIterations=1000
RunIterations=2000
Chains=5

[Scenario_Script]
BurnInIterations=1000
RunIterations=2000
Chains=1
```

The *required fields* section defines which column names must be present in the import data files.

These are the fields that the importer will validate is present.

```
[RequiredFields]
FoodImportExport=Country_From;Country_To;Food_Source;Amount
Prevalence=Country;Genus;Species;Food_Source;Units_Tested;Units_Positive
HumanCase=Country;Genus;Species;Incidence;Travel;Outbreak
```

Following field names can be used for FoodImportExport:

- Year
- Country_From
- Country_To
- Food_Source
- Amount

Following field names can be used for Prevalence:

- Year
- Country
- Genus
- Species
- Food_Source
- Units_Tested
- Units_Positive

Following field names can be used for HumanCase:

- Year
- Country

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

- Genus
- Species
- Incidence
- Travel
- Travel_Yes
- Travel_No
- Travel_Unknown
- Outbreak

For the single member state model the fields year, travel_yes, travel_no and travel_unknown is used

For multiple member state model the fields country, country_from, country_to and Travel are used.

Dimension Ranges

In order for the WinBUGS code to determine the ranges of the different dimensions, a set of Dimension Count variables are available:

ReportingCountryCount – The number of selected reporting countries
SourceCountryCount – The number of selected source countries
FoodSourceCount – The number of selected food sources
SeroVarCount – The number of selected serovars
YearCount – The number of selected years

So if e.g. three food sources have been selected, FoodSourceCount will be set to 3 and inserted into the WinBUGS data file.

These ranges can be used throughout the *Variable Definitions* configuration and in the WinBUGS code itself.

Dimension Minimums

In order to have a proper execution of the model, it is possible to set a minimum number of values that must be selected before running the model:

```
[DimensionMinimums]
FoodSources=3
ReportingCountries=5
SourceCountries=5
Years=2
Subtypes=10
```

Data

The *data* section defines the names of the variables and their dimensions. EFSA_SAM makes output data to WinBUGSS based on these settings.

```
[Data]
Incidence=n[ReportingCountryCount,SeroVarCount]
Travel=yt[ReportingCountryCount,SeroVarCount]
Outbreak=outb[ReportingCountryCount,SeroVarCount]
OutbreakSum=outbreak[ReportingCountryCount]
ImportExport=m[ReportingCountryCount,SourceCountryCount,FoodSourceCount]
Prevalence=p[SourceCountryCount,FoodSourceCount,SeroVarCount]
```

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Underreporting=u[ReportingCountryCount], consisting of:

- MeanUnderReporting=m.u[ReportingCountryCount]
- StandardDeviationUnderReporting=s.u[ReportingCountryCount]

Distributions

The *Distributions* section defines the distributions and their dimensions

```
[Distributions]
a=ReportingCountryCount,FoodSourceCount
q=SeroVarCount
u=ReportingCountryCount

[Distribution_a]
Grouping=FoodSourceCount
Default_Range=0,100
Visible=True
FloatingPoint=False
UseMaxRangeArray=True
Transpose=True

[Distribution_q]
Grouping=
Default_Range=0,100
NA_Elements=1,SeroVarCount
Visible=True
FloatingPoint=False
UseMaxRangeArray=True
Transpose=False

[Distribution_u]
Grouping=
Default_Range=-2,5
Visible=True
FloatingPoint=True
UseMaxRangeArray=False
Transpose=False
```

The *Grouping* defines whether groups of initial values are needed.
In this example, distribution *a* is defined by the dimensions ReportingCountryCount and FoodSourceCount.

If there are 24 reporting countries and 4 food sources, there will be 96 elements. If no grouping is selected, it will be a single dimension array containing 96 elements.

If (like in the example above) a grouping is applied (this case Grouping=FoodSourceCount)
a will be defined as a two-dimensional array of 4 rows by 24 columns.
If Transpose=True, then the array is transposed and will now be 24 rows by 4 columns.

If Floating point is false, the data elements will be integer numbers, otherwise they will be floating points.

If UseMaxRangeArray=True then an additional array will be created with the values of Maximum (upper range setting) of the distribution. It may look like this:

```
# Distribution Ranges for a
a_MaxRange=c(100,100,100,100),
```

And in the code it may be used like this:

```
# Food-source-dependent factor (a): two-dimensional and uniform priors
for (c in 1:ReportingCountryCount)
{
        for (j in 1:FoodSourceCount)
        {
                a[c,j] ~ dunif(0,a_MaxRange[j])
        }
}
```

Kernel Density Plot

The *KernelDensityPlot* section simply tells which distributions density plots are made for.

```
[KernelDensityPlot]
a=Coda
q=Coda
```

WinBUGS_Defaults

The *WinBUGS_Defaults* section defines the default values for the number of chains and for the number of Burn In and Run iterations.

```
[WinBUGS_Defaults]
BurnInIterations=10000
RunIterations=30000
Chains=3
```

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Output_Variables

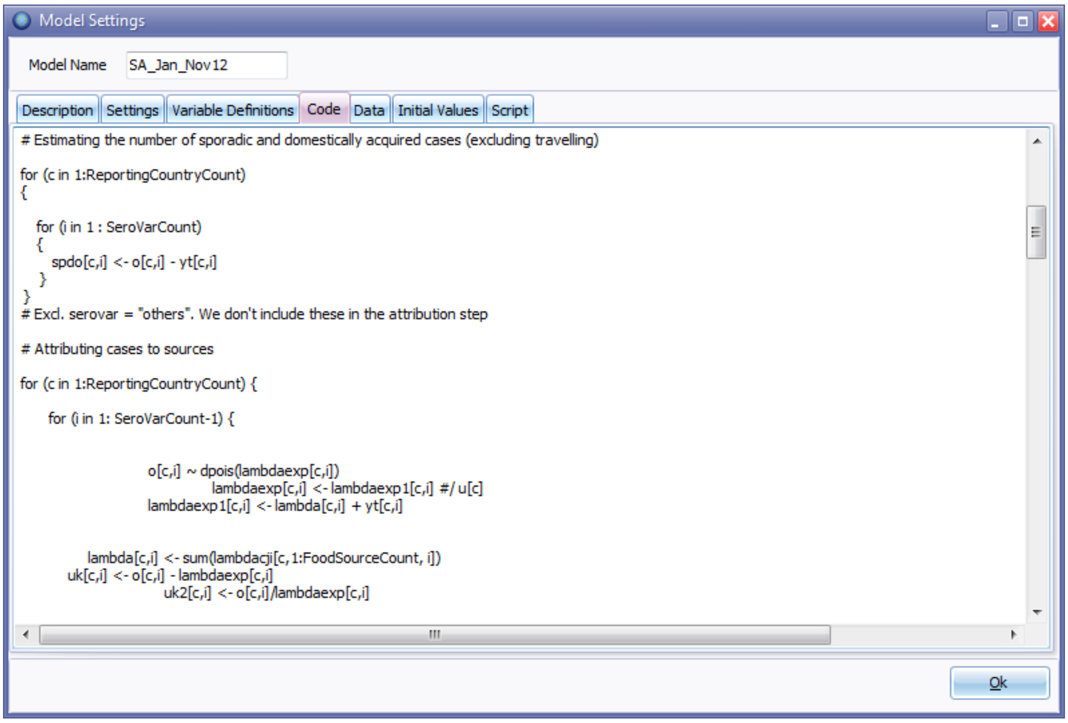
The *Output_Variables* section, defines which variables from the code are available to EFSA_SAM and that subsequently be produced by WinBUGS when the model is run.

```
[Output_Variables]
serotype[FoodSourceCount,SeroVarCount]=Number of estimated cases in EU per source and serotype;EU;Baseline
serotypescenario[FoodSourceCount,SeroVarCount]=Number of estimated cases in EU per source and serotype;EU;Scenario
source[ReportingCountryCount,FoodSourceCount]=Number of estimated cases per source and reporting MS;MS;Baseline
sourcescenario[ReportingCountryCount,FoodSourceCount]=Number of estimated cases per source and reporting MS;MS;Scenario
source2[SourceCountryCount,FoodSourceCount]=Number of estimated cases per source and MS of origin;MS;Baseline
source2scenario[SourceCountryCount,FoodSourceCount]=Number of estimated cases per source and MS of origin;MS;Scenario
truecasescji[ReportingCountryCount,FoodSourceCount,SeroVarCount]=Number of cases pr. MS, food source and serotype;MS;Baseline
truecasesscenariocji[ReportingCountryCount,FoodSourceCount,SeroVarCount]=number of cases pr. MS, food source and serotype;MS;Scenario
travel[ReportingCountryCount]=Number of estimated travel-related cases per MS;MS;Baseline+Scenario
unknown[ReportingCountryCount]=Number of estimated cases with unknown source per MS;MS;Baseline+Scenario
total[ReportingCountryCount]=Number of estimated total cases per MS;MS;Baseline
totalscenario[ReportingCountryCount]=Number of estimated total cases per MS;MS;Scenario
GoodFit[ReportingCountryCount]=Goodness of fit ratio: reported number of cases per MS divided by the estimated number of cases per MS per serovar.;MS;Baseline+Scenario
totalEU=Number of estimated total cases in EU;EU;Baseline
unknownEU=Number of estimated cases with unknown source in EU;EU;Baseline+Scenario
travelEU=Number of estimated travel-related cases in EU;EU;Baseline+Scenario
unktravEU=Sum of unknownEU and travel EU;EU;Baseline+Scenario
totalEUscenario=Number of estimated total cases in EU;EU;Scenario
difftotalEU=totalEU – totalEUscen1;EU;Scenario
sourceC[FoodSourceCount]=Number of estimated cases per source in EU;EU;Baseline
percsourceC[FoodSourceCount]=Percentage of estimated cases per source in EU;EU;Baseline
sourceCscenario[FoodSourceCount]=Number of estimated cases per source in EU;EU;Scenario
percsourceCscenario[FoodSourceCount]=Percentage of estimated cases per source in EU;EU;Scenario
diffsourceC[FoodSourceCount]=sourceC – sourceCscen1;EU;Scenario
diffpercsourceC[FoodSourceCount]=percsourceC – percsourceCscen1;EU;Scenario
percdiffsourceC[FoodSourceCount]=Percentage difference between baseline and scenario (diffsourceC=100/sourceC);EU;Scenario
a[ReportingCountryCount,FoodSourceCount]=Food-source related factor per source and MS;EU+MS;Baseline+Scenario
q[SeroVarCount]=Serovar related factor. One per serovar;EU+MS;Baseline+Scenario
uf[ReportingCountryCount] =Underreporting factor. One per MS;EU+MS;Baseline+Scenario
```

10.3. Model code

The model code is standard WinBUGS code. In order for the code to work along with the settings that EFSA_SAM produces, some variables can be used for controlling the loops.

The variables will be the count of Reporting Countries, Food Sources, Serovars, Source Countries, ensuring that the loops will run within the limits of the data arrays, also produced by EFSA_SAM, based on the selections the user makes during analysis setup.



Model Name SA_Jan_Nov12

```
# Estimating the number of sporadic and domestically acquired cases (excluding travelling)

for (c in 1:ReportingCountryCount)
{

   for (i in 1 : SeroVarCount)
   {
      spdo[c,i] <- o[c,i] - yt[c,i]
   }
}
# Excl. serovar = "others". We don't include these in the attribution step

# Attributing cases to sources

for (c in 1:ReportingCountryCount) {

     for (i in 1: SeroVarCount-1) {


                        o[c,i] ~ dpois(lambdaexp[c,i])
                                lambdaexp[c,i] <- lambdaexp1[c,i] #/ u[c]
                        lambdaexp1[c,i] <- lambda[c,i] + yt[c,i]


             lambda[c,i] <- sum(lambdacji[c,1:FoodSourceCount, i])
          uk[c,i] <- o[c,i] - lambdaexp[c,i]
                          uk2[c,i] <- o[c,i]/lambdaexp[c,i]
```

It is highly recommended to describe/document all variables used in the code. This can be done, either as comments in the code or in the description section of the model.

Also meaningful names for the variables are highly recommended.

This will make it easier for other people to understand what is going on in the inner workings of the model.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

B. DATA TEMPLATES FOR THE EFSA_SAM MODEL

Table B1. Example of data template for reported cases of human salmonellosis.

Year	Country	Genus	Species	Serovar	Total_cases	Travel_cases	Domestic_cases	Unknown_travel	Outbreak_cases
2010	MS1	Salmonella	S.enterica	AGONA	37	9	21	7	0
2010	MS1	Salmonella	S.enterica	ANATUM	13	4	7	2	0
2010	MS1	Salmonella	S.enterica	BOVISMORBIFICANS	14	0	10	4	0
2010	MS1	Salmonella	S.enterica	BRAENDERUP	35	17	13	5	0
2010	MS1	Salmonella	S.enterica	BRANDENBURG	9	0	7	2	0
2010	MS1	Salmonella	S.enterica	BREDENEY	8	0	6	2	0
2010	MS1	Salmonella	S.enterica	DERBY	10	1	7	2	0
2010	MS1	Salmonella	S.enterica	ENTERITIDIS	6102	673	4072	1357	217
2010	MS1	Salmonella	S.enterica	HADAR	62	22	30	10	0
2010	MS1	Salmonella	S.enterica	HEIDELBERG	16	2	10	4	0
2010	MS1	Salmonella	S.enterica	INFANTIS	107	5	76	26	0
2010	MS1	Salmonella	S.enterica	KENTUCKY	22	11	8	3	0
2010	MS1	Salmonella	S.enterica	KOTTBUS	17	5	9	3	0
2010	MS1	Salmonella	S.enterica	LIVINGSTONE	10	0	7	3	0
etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Table B2. Example of data template for reported animal food data.

Year	Country	Genus	Species	Serovar	Food_Source	Units_Tested	Units positive
2008	MS1	Salmonella	S.enterica	INFANTIS	Broilers	408	1
2008	MS1	Salmonella	S.enterica	KENTUCKY	Broilers	408	1
2008	MS1	Salmonella	S.enterica	TYPHIMURIUM	Broilers	408	1
2008	MS1	Salmonella	S.enterica	ENTERITIDIS	Broilers	408	2
2008	MS1	Salmonella	S.enterica	MONTEVIDEO	Broilers	408	4
2009	MS1	Salmonella	S.enterica	HADAR	Broilers	380	1
2009	MS1	Salmonella	S.enterica	INFANTIS	Broilers	380	1
2009	MS1	Salmonella	S.enterica	KENTUCKY	Broilers	380	2
2009	MS1	Salmonella	S.enterica	TYPHIMURIUM	Broilers	380	4
2009	MS1	Salmonella	S.enterica	ENTERITIDIS	Broilers	380	4
2009	MS1	Salmonella	S.enterica	MONTEVIDEO	Broilers	380	6
2009	MS1	Salmonella	S.enterica	INFANTIS	Broilers	380	8
2010	MS1	Salmonella	S.enterica	TYPHIMURIUM	Broilers	450	13
etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Table B3. Example of data template for animal food production and trade data.

Year	Country_From	Country_To	Food_Source	Tons
2010	AT	AT	Broilers	239,681
2010	AT	BE	Broilers	60
2010	AT	CY	Broilers	0
2010	AT	CZ	Broilers	1,336
2010	AT	DE	Broilers	42,965
2010	AT	DK	Broilers	6
2010	AT	EE	Broilers	54
2010	AT	ES	Broilers	23
2010	AT	FI	Broilers	0
2010	AT	FR	Broilers	103
2010	AT	GR	Broilers	102
etc.	etc.	etc.	etc.	etc.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

C. Data templates for the Single Member State model

Table C1. Example of data template for reported cases of human salmonellosis.

Year	Genus	Species	Serotype	Phage type	Resistance profile	Incidence	Outbreak	Travel_yes	Travel_no[a]	Travel_uk
2009	Salmonella	Enterica	Enteritidis	PT 21		7	0	4	1	2
2009	Salmonella	Enterica	Enteritidis	PT 4		6	0	1	4	1
2009	Salmonella	Enterica	Enteritidis	PT 8		8	0	0	6	2
2009	Salmonella	Enterica	Typhimurium	DT 104	RRRRSSSSS	9	0	0	7	2
2009	Salmonella	Enterica	Typhimurium	DT 120	RSRRSSSSS	20	11	2	16	2
2009	Salmonella	Enterica	Typhimurium	DT 120	RSRSSSSSS	6	1	0	5	1
2009	Salmonella	Enterica	Typhimurium	DT 193	RSRRSSSSS	41	5	17	20	4
2009	Salmonella	Enterica	Typhimurium	NT	RSRSSSSSS	1	0	0	1	0
2009	Salmonella	Enterica	Typhimurium	U302	RRRRSSSSS	2	0	1	0	1
2009	Salmonella	Enterica	Typhimurium	DT 12	SSSSSSSSS	10	0	2	8	0
2009	Salmonella	Enterica	Typhimurium	DT 120	SSSSSSSSS	3	0	0	2	1
2009	Salmonella	Enterica	Typhimurium	DT 17	SSSSSSSSS	5	0	0	4	1
2009	Salmonella	Enterica	Typhimurium	DT 193	SSSSSSSSS	1	0	0	1	0
2009	Salmonella	Enterica	Typhimurium	DT 41	SSSSSSSSS	0	0	0	0	0
2009	Salmonella	Enterica	Typhimurium	DT 66	SSSSSSSSS	0	0	0	0	0
2009	Salmonella	Enterica	Typhimurium	DT 8	SSSSSSSSS	4	0	0	2	2
2009	Salmonella	Enterica	Typhimurium	NT	SSSSSSSSS	21	0	2	14	5
2009	Salmonella	Enterica	Typhimurium	U312	SSSSSSSSS	3	0	0	1	2
2009	Salmonella	Enterica	Typhimurium	DT 193		1	1	1	0	0
2009	Salmonella	Enterica	Agona			13	0	9	2	2
2009	Salmonella	Enterica	Anatum			3	0	0	1	2
2009	Salmonella	Enterica	Brandenburg			3	0	1	1	1
etc.	etc.	etc.	etc.		etc.	etc.	etc.	etc.	etc.	etc.

a) If the data do not distinguish between no to travel (travel_no) and unknown travel history (travel_uk), these cases should be reported as travel_no.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Table C2. Example of data template for reported animal food data.

Year	Genus	Species	Serotype	Phage type	Resistance profile	Food_Source	Units_Tested	Units_positive
2009	Salmonella	Enterica	Enteritidis	PT 21		BEEF	2267	0
2009	Salmonella	Enterica	Enteritidis	PT 4		BEEF	2267	0
2009	Salmonella	Enterica	Enteritidis	PT 8		BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 104	RRRRSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 120	RSRRSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 120	RSRSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 193	RSRRSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	NT	RSRSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	U302	RRRRSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 12	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 120	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 17	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 193	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 41	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 66	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 8	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	NT	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	U312	SSSSSSSSS	BEEF	2267	0
2009	Salmonella	Enterica	Typhimurium	DT 193		BEEF	2267	0
2009	Salmonella	Enterica	Agona			BEEF	2267	0
2009	Salmonella	Enterica	Anatum			BEEF	2267	0
2009	Salmonella	Enterica	Brandenburg			BEEF	2267	0
2009	Salmonella	Enterica	Derby			BEEF	2267	0
2009	Salmonella	Enterica	Dublin			BEEF	2267	2
etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Table C3. Example of data template for animal food production data.

Year	Food_Source	Amount
2009	PORK	90928171.5
2009	BEEF	46099552
2009	LAYERS	31100000
2009	BROILERS	26580904.5
2009	DUCK	403698
2009	IMP_PORK	43651806
2009	IMP_BEEF	51305993.5
2009	IMP_CHICK	21333407.5
2009	IMP_DUCK	4370019
2009	IMP_TURKEY	2437732.5
2010	PORK	90930660.5
2010	BEEF	46101118
2010	LAYERS	31101230
2010	BROILERS	26582393.5
2010	DUCK	403818
2010	IMP_PORK	43657206
2010	IMP_BEEF	51311262.5
2010	IMP_CHICK	21336663.5
2010	IMP_DUCK	4370143
2010	IMP_TURKEY	2437901.5
2011	PORK	90930330.5
etc.	etc.	etc.

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

D. Variable definitions and model codes for the TT-SAM and Single Member State model

D1 – Variable definitions for TT-SAM with underreporting

```
[RequiredFields]
FoodImportExport=Country_From;Country_To;Food_Source;Amount
Prevalence=Country;Genus;Species;Food_Source;Units_Tested;Units_Positive
HumanCase=Country;Genus;Species;Incidence;Travel;Outbreak

[DimensionMinimums]
FoodSources=3
ReportingCountries=5
SourceCountries=5
Subtypes=5

[Distributions]
a=ReportingCountryCount,FoodSourceCount
q=SeroVarCount
u=ReportingCountryCount

[Data]
Incidence=n[ReportingCountryCount,SeroVarCount]
Travel=yt[ReportingCountryCount,SeroVarCount]
Outbreak=outb[ReportingCountryCount,SeroVarCount]
OutbreakSum=outbreak[ReportingCountryCount]
ImportExport=m[ReportingCountryCount,SourceCountryCount,FoodSourceCount]
Prevalence=p[SourceCountryCount,FoodSourceCount,SeroVarCount]
UnderReporting=u[ReportingCountryCount]

[Distribution_a]
Grouping=FoodSourceCount
Default_Range=0,100
Visible:=True
FloatingPoint=False
UseMaxRangeArray=True
Transpose=True

[Distribution_q]
Grouping=
Default_Range=0,100
NA_Elements=1,SeroVarCount
Visible:=True
FloatingPoint=False
UseMaxRangeArray=True
Transpose=False

[Distribution_u]
Grouping=
Default_Range=-2,5
Visible:=True
FloatingPoint=True
UseMaxRangeArray=False
Transpose=False

[KernelDensityPlot]
a=Coda
q=Coda

[WinBUGS_Defaults]
BurnInIterations=10000
```

```
RunIterations=30000
Chains=3

[Output_Variables]
serotype[FoodSourceCount,SeroVarCount]=Number of estimated cases in EU per source and serotype;EU;Baseline
serotypescenario[FoodSourceCount,SeroVarCount]=Number of estimated cases in EU per source and serotype;EU;Scenario
source[ReportingCountryCount,FoodSourceCount]=Number of estimated cases per source and reporting MS;MS;Baseline
sourcescenario[ReportingCountryCount,FoodSourceCount]=Number of estimated cases per source and reporting MS;MS;Scenario
source2[SourceCountryCount,FoodSourceCount]=Number of estimated cases per source and MS of origin;MS;Baseline
source2scenario[SourceCountryCount,FoodSourceCount]=Number of estimated cases per source and MS of origin;MS;Scenario
truecasescji[ReportingCountryCount,FoodSourceCount,SeroVarCount]=Number of cases pr. MS, food source and serotype;MS;Baseline
truecasesscenariocji[ReportingCountryCount,FoodSourceCount,SeroVarCount]=number of cases pr. MS, food source and serotype;MS;Scenario
travel[ReportingCountryCount]=Number of estimated travel-related cases per MS;MS;Baseline+Scenario
unknown[ReportingCountryCount]=Number of estimated cases with unknown source per MS;MS;Baseline+Scenario
total[ReportingCountryCount]=Number of estimated total cases per MS;MS;Baseline
totalscenario[ReportingCountryCount]=Number of estimated total cases per MS;MS;Scenario
GoodFit[ReportingCountryCount]=Goodness of fit ratio: reported number of cases per MS divided by the estimated number of cases per MS per serovar.;MS;Baseline+Scenario
totalEU=Number of estimated total cases in EU;EU;Baseline
unknownEU=Number of estimated cases with unknown source in EU;EU;Baseline+Scenario
travelEU=Number of estimated travel-related cases in EU;EU;Baseline+Scenario
unktravEU=Sum of unknownEU and travel EU;EU;Baseline+Scenario
totalEUscenario=Number of estimated total cases in EU;EU;Scenario
difftotalEU=totalEU – totalEUscen1;EU;Scenario
sourceC[FoodSourceCount]=Number of estimated cases per source in EU;EU;Baseline
percsourceC[FoodSourceCount]=Percentage of estimated cases per source in EU;EU;Baseline
sourceCscenario[FoodSourceCount]=Number of estimated cases per source in EU;EU;Scenario
percsourceCscenario[FoodSourceCount]=Percentage of estimated cases per source in EU;EU;Scenario
diffsourceC[FoodSourceCount]=sourceC – sourceCscen1;EU;Scenario
diffpercsourceC[FoodSourceCount]=percsourceC – percsourceCscen1;EU;Scenario
percdiffsourceC[FoodSourceCount]=Percentage difference between baseline and scenario (diffsourceC=100/sourceC);EU;Scenario
a[ReportingCountryCount,FoodSourceCount]=Food-source related factor per source and MS;EU+MS;Baseline+Scenario
q[SeroVarCount]=Serovar related factor. One per serovar;EU+MS;Baseline+Scenario
uf[ReportingCountryCount] =Underreporting factor. One per MS;EU+MS;Baseline+Scenario
```

D2 – WinBUGS code for TT-SAM with underreporting

```
# Model

{

# SALMONELLA including others

# Estimating the number of sporadic and domestically acquired cases (excluding travel and outbreak cases)

for (c in 1:ReportingCountryCount)
{

    for (i in 1:SeroVarCount)
        {

                ob1[c,i] <- outb[c,i] - 1
                ob2[c,i] <- max(ob1[c,i],0)
                o[c,i] <- n[c,i] - ob2[c,i]
                spdo[c,i] <- o[c,i] - yt[c,i]
        }
}

# Excl. "others"

# Attributing cases to sources

for (c in 1:ReportingCountryCount)
{
        for (i in 1:SeroVarCount-1)
        {
                o[c,i] ~ dpois(lambdaexp[c,i])
                lambdaexp[c,i] <- lambda[c,i] + yt[c,i]
                 lambda[c,i] <- sum(lambdacji[c,1:FoodSourceCount, i])
                 uk[c,i] <- o[c,i] - lambdaexp[c,i]

                  for (j in 1:FoodSourceCount)
                 {
                        for (k in 1:SourceCountryCount)
                        {
                                lambdackji[c,k,j,i] <- p[k,j,i]*m[c,k,j]*a[c,j]*q[i]
                                truecasesckji[c,k,j,i]<- lambdackji[c,k,j,i]*uf[c]

                        }
                        lambdacji[c,j,i] <- sum(lambdackji[c,1:SourceCountryCount,j,i])
                        truecasescji[c,j,i] <- sum(truecasesckji[c,1:SourceCountryCount,j,i])

                 }
        }

}


for (c in 1:ReportingCountryCount)
{
        for (i in 1: SeroVarCount-1)
        {
                for (j in 1:FoodSourceCount)
                {
                        for (k in 1:SourceCountryCount)
                        {
                                casescenariockji[c,k,j,i] <- p_scen[k,j,i]*m[c,k,j]*a[c,j]*q[i]
                                truecasesscenariockji[c,k,j,i]<- casescenariockji[c,k,j,i]*uf[c]
```

```
            }
            truecasesscenariocji[c,j,i] <- sum(truecasesscenariockji[c,1:SourceCountryCount,j,i])
        }
    }
}


# summing over k by c

for (j in 1:FoodSourceCount)
{
    for (k in 1:SourceCountryCount)
    {
        for (c in 1:ReportingCountryCount)
        {
            truecasesorigin[c,k,j] <- sum(truecasesckji[c,k,j,1:SeroVarCount-1])
            truecasesoriginscenario[c,k,j] <- sum(truecasesscenariockji[c,k,j,1:SeroVarCount-1])
        }
    }
}

# OUTPUT

# serovar per source

for (i in 1: SeroVarCount-1)
{
    for (j in 1:FoodSourceCount)
    {
        serotype[j,i] <- sum(truecasescji[1:ReportingCountryCount,j,i])
        serotypescenario[j,i] <- sum(truecasesscenariocji[1:ReportingCountryCount,j,i])
    }
}

for (c in 1:ReportingCountryCount)
{
    for(i in 1:SeroVarCount-1)
    {
        truecasesci[c,i] <- sum(truecasescji[c,1:FoodSourceCount,i])
        truecasesscenarioci[c,i] <- sum(truecasesscenariocji[c,1:FoodSourceCount,i])
    }

    for (j in 1:FoodSourceCount)
    {

        source[c,j] <- sum(truecasescji[c,j,1:SeroVarCount-1])
        sourcescenario[c,j] <- sum(truecasesscenariocji[c,j,1:SeroVarCount-1])
    }
}


for (k in 1:SourceCountryCount)
{
    for (j in 1:FoodSourceCount)
    {
        source2[k,j] <- sum(truecasesorigin[1:ReportingCountryCount,k,j])
        source2scenario[k,j] <- sum(truecasesoriginscenario[1:ReportingCountryCount,k,j])
    }
}


for (c in 1:ReportingCountryCount)
{
        travel[c]<- sum(yt[c,])*uf[c]
        unknown1[c] <- spdo[c,SeroVarCount] + sum(uk[c,1:SeroVarCount-1])
        unknown[c] <- unknown1[c]*uf[c]

        total[c]<- travel[c] + unknown[c] + sum(source[c,1:FoodSourceCount]) +  outbreak[c]
        totalscenario[c]<- travel[c] + unknown[c] + sum(sourcescenario[c,1:FoodSourceCount]) +  outbreak[c]

        GoodFit[c] <- o[c,1:SeroVarCount-1]/lambdaexp[c,1:SeroVarCount-1]  # for evaluating goodness of fit; excluding other serovars
}

totalEU <- sum(total[1:ReportingCountryCount])
unknownEU <- sum(unknown[1:ReportingCountryCount])
travelEU<- sum(travel[1:ReportingCountryCount])
unktravEU <- unknownEU + travelEU

totalEUscenario <- sum(totalscenario[1:ReportingCountryCount])
difftotalEU <- totalEU - totalEUscenario


for (j in 1:FoodSourceCount)
{
        sourceC[j] <- sum(source[1:ReportingCountryCount,j])
        percsourceC[j] <- sourceC[j]*100/totalEU

        sourceCscenario[j] <- sum(sourcescenario[1:ReportingCountryCount,j])
        percsourceCscenario[j] <- sourceCscenario[j]*100/totalEUscenario

        diffsourceC[j] <- sourceC[j] - sourceCscenario[j]
        diffpercsourceC[j] <- percsourceC[j] - percsourceCscenario[j]
        percdiffsourceC[j] <- diffsourceC[j]*100/sourceC[j]
}

# DEFINITION OF PRIORS

# Food-source-dependent factor (a): two-dimensional and uniform priors

for (c in 1:ReportingCountryCount)
{
        for (j in 1:FoodSourceCount)
        {
                a[c,j] ~ dunif(0,a_MaxRange[j])
        }
}

# Subtype-dependent factor (q): uni-dimensional and uniform priors

q[1] <- 1  # q for SE set to 1; all others will be estimated relative to this serovar
q[SeroVarCount] <- 1 # not used for estimating cases; value insignificant

for (i in 2:SeroVarCount-1)
{
        q[i] ~  dunif(0,q_MaxRange[1])
}

# Underreporting factor, defined as a lognormal distribution: replace mu and tau for mean and sdev

for (c in 1:ReportingCountryCount)
{
    uf[c] <- exp(u[c])
    u[c] ~ dnorm(m.u[c],pr.u[c])
    pr.u[c] <- pow(s.u[c],-2)
}
```

```
} # End of model
```

D3 – Variable definitions for TT-SAM without (wo) underreporting

```
[RequiredFields]
FoodImportExport=Country_From;Country_To;Food_Source;Amount
Prevalence=Country;Genus;Species;Food_Source;Units_Tested;Units_Positive
HumanCase=Country;Genus;Species;Incidence;Travel;Outbreak

[DimensionMinimums]
FoodSources=3
ReportingCountries=5
SourceCountries=5
Subtypes=5

[Distributions]
a=ReportingCountryCount,FoodSourceCount
q=SeroVarCount

[Data]
Incidence=n[ReportingCountryCount,SeroVarCount]
Travel=yt[ReportingCountryCount,SeroVarCount]
Outbreak=outb[ReportingCountryCount,SeroVarCount]
OutbreakSum=outbreak[ReportingCountryCount]
ImportExport=m[ReportingCountryCount,SourceCountryCount,FoodSourceCount]
Prevalence=p[SourceCountryCount,FoodSourceCount,SeroVarCount]

[Distribution_a]
Grouping=FoodSourceCount
Default_Range=0,100
Visible:=True
FloatingPoint=False
UseMaxRangeArray=True
Transpose=True

[Distribution_q]
Grouping=
Default_Range=0,100
NA_Elements=1,SeroVarCount
Visible:=True
FloatingPoint=False
UseMaxRangeArray=True
Transpose=False

[KernelDensityPlot]
a=Coda
q=Coda

[WinBUGS_Defaults]
BurnInIterations=10000
RunIterations=30000
Chains=3


[Output_Variables]
serotype[FoodSourceCount,SeroVarCount]=Number of estimated cases in EU per source and serotype;EU;Baseline
serotypescenario[FoodSourceCount,SeroVarCount]=Number of estimated cases in EU per source and serotype;EU;Scenario
source[ReportingCountryCount,FoodSourceCount]=Number of estimated cases per source and reporting MS;MS;Baseline
sourcescenario[ReportingCountryCount,FoodSourceCount]=Number of estimated cases per source and reporting MS;MS;Scenario
source2[SourceCountryCount,FoodSourceCount]=Number of estimated cases per source and MS of origin;MS;Baseline
source2scenario[SourceCountryCount,FoodSourceCount]=Number of estimated cases per source and MS of origin;MS;Scenario
truecasescji[ReportingCountryCount,FoodSourceCount,SeroVarCount]=Number of cases pr. MS, food source and serotype;MS;Baseline
```

truecasesscenariocji[ReportingCountryCount,FoodSourceCount,SeroVarCount]=number of cases pr. MS, food source and serotype;MS;Scenario

travel[ReportingCountryCount]=Number of estimated travel-related cases per MS;MS;Baseline+Scenario

unknown[ReportingCountryCount]=Number of estimated cases with unknown source per MS;MS;Baseline+Scenario

total[ReportingCountryCount]=Number of estimated total cases per MS;MS;Baseline

totalscenario[ReportingCountryCount]=Number of estimated total cases per MS;MS;Scenario

GoodFit[ReportingCountryCount]=Goodness of fit ratio: reported number of cases per MS divided by the estimated number of cases per MS per serovar.;MS;Baseline+Scenario

totalEU=Number of estimated total cases in EU;EU;Baseline

unknownEU=Number of estimated cases with unknown source in EU;EU;Baseline+Scenario

travelEU=Number of estimated travel-related cases in EU;EU;Baseline+Scenario

unktravEU=Sum of unknownEU and travel EU;EU;Baseline+Scenario

totalEUscenario=Number of estimated total cases in EU;EU;Scenario

difftotalEU=totalEU – totalEUscen1;EU;Scenario

sourceC[FoodSourceCount]=Number of estimated cases per source in EU;EU;Baseline

percsourceC[FoodSourceCount]=Percentage of estimated cases per source in EU;EU;Baseline

sourceCscenario[FoodSourceCount]=Number of estimated cases per source in EU;EU;Scenario

percsourceCscenario[FoodSourceCount]=Percentage of estimated cases per source in EU;EU;Scenario

diffsourceC[FoodSourceCount]=sourceC – sourceCscen1;EU;Scenario

diffpercsourceC[FoodSourceCount]=percsourceC – percsourceCscen1;EU;Scenario

percdiffsourceC[FoodSourceCount]=Percentage difference between baseline and scenario (diffsourceC=100/sourceC);EU;Scenario

a[ReportingCountryCount,FoodSourceCount]=Food-source related factor per source and MS;EU+MS;Baseline+Scenario

q[SeroVarCount]=Serovar related factor. One per serovar;EU+MS;Baseline+Scenario

uf[ReportingCountryCount] =Underreporting factor. One per MS;EU+MS;Baseline+Scenario

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

D4 – WinBUGS code for TT-SAM without (wo) underreporting

```
# Model

{

# SALMONELLA including others

# Estimating the number of sporadic and domestically acquired cases (excluding travel and outbreak cases)

for (c in 1:ReportingCountryCount)
{
    for (i in 1:SeroVarCount)
        {
                ob1[c,i] <- outb[c,i] - 1
                ob2[c,i] <- max(ob1[c,i],0)
                o[c,i] <- n[c,i] - ob2[c,i]
                spdo[c,i] <- o[c,i] - yt[c,i]
        }
}

# Excl. "others"
# Attributing cases to sources

for (c in 1:ReportingCountryCount)
{
        for (i in 1:SeroVarCount-1)
        {
                o[c,i] ~ dpois(lambdaexp[c,i])
                lambdaexp[c,i] <- lambda[c,i] + yt[c,i]
                 lambda[c,i] <- sum(lambdacji[c,1:FoodSourceCount, i])
                 uk[c,i] <- o[c,i] - lambdaexp[c,i]

                  for (j in 1:FoodSourceCount)
                 {
                        for (k in 1:SourceCountryCount)
                        {
                                lambdackji[c,k,j,i] <- p[k,j,i]*m[c,k,j]*a[c,j]*q[i]
                                truecasesckji[c,k,j,i]<- lambdackji[c,k,j,i]

                        }
                        lambdacji[c,j,i] <- sum(lambdackji[c,1:SourceCountryCount,j,i])
                        truecasescji[c,j,i] <- sum(truecasesckji[c,1:SourceCountryCount,j,i])

                 }
        }

}


for (c in 1:ReportingCountryCount)
{
        for (i in 1: SeroVarCount-1)
        {
                for (j in 1:FoodSourceCount)
                {
                        for (k in 1:SourceCountryCount)
                        {
                                casescenariockji[c,k,j,i] <- p_scen[k,j,i]*m[c,k,j]*a[c,j]*q[i]
                                truecasesscenariockji[c,k,j,i]<- casescenariockji[c,k,j,i]
                        }
                        truecasesscenariocji[c,j,i] <- sum(truecasesscenariockji[c,1:SourceCountryCount,j,i])
                }
```

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

```
        }
}


# summing over k by c

for (j in 1:FoodSourceCount)
{
        for (k in 1:SourceCountryCount)
        {
                for (c in 1:ReportingCountryCount)
                {
                        truecasesorigin[c,k,j] <- sum(truecasesckji[c,k,j,1:SeroVarCount-1])
                        truecasesoriginscenario[c,k,j] <- sum(truecasesscenariockji[c,k,j,1:SeroVarCount-1])
                }
        }
}

# OUTPUT

# serovar per source

for (i in 1: SeroVarCount-1)
{
        for (j in 1:FoodSourceCount)
        {
                serotype[j,i] <- sum(truecasescji[1:ReportingCountryCount,j,i])
                serotypescenario[j,i] <- sum(truecasesscenariocji[1:ReportingCountryCount,j,i])
        }
}

for (c in 1:ReportingCountryCount)
{
        for(i in 1:SeroVarCount-1)
        {
                truecasesci[c,i] <- sum(truecasescji[c,1:FoodSourceCount,i])
                truecasesscenarioci[c,i] <- sum(truecasesscenariocji[c,1:FoodSourceCount,i])
        }

        for (j in 1:FoodSourceCount)
        {

                source[c,j] <- sum(truecasescji[c,j,1:SeroVarCount-1])
                sourcescenario[c,j] <- sum(truecasesscenariocji[c,j,1:SeroVarCount-1])
        }
}


for (k in 1:SourceCountryCount)
{
        for (j in 1:FoodSourceCount)
        {
                source2[k,j] <- sum(truecasesorigin[1:ReportingCountryCount,k,j])
                source2scenario[k,j] <- sum(truecasesoriginscenario[1:ReportingCountryCount,k,j])
        }
}


for (c in 1:ReportingCountryCount)
{
        travel[c]<- sum(yt[c,])
        unknown1[c] <- spdo[c,SeroVarCount] + sum(uk[c,1:SeroVarCount-1])
        unknown[c] <- unknown1[c]
```

```
    total[c]<- travel[c] + unknown[c] + sum(source[c,1:FoodSourceCount]) +  outbreak[c]
    totalscenario[c]<- travel[c] + unknown[c] + sum(sourcescenario[c,1:FoodSourceCount]) +  outbreak[c]

    GoodFit[c] <- o[c,1:SeroVarCount-1]/lambdaexp[c,1:SeroVarCount-1]  # for evaluating goodness of fit; excluding other serovars
}

totalEU <- sum(total[1:ReportingCountryCount])
unknownEU <- sum(unknown[1:ReportingCountryCount])
travelEU<- sum(travel[1:ReportingCountryCount])
unktravEU <- unknownEU + travelEU

totalEUscenario <- sum(totalscenario[1:ReportingCountryCount])
difftotalEU <- totalEU - totalEUscenario


for (j in 1:FoodSourceCount)
{
    sourceC[j] <- sum(source[1:ReportingCountryCount,j])
    percsourceC[j] <- sourceC[j]*100/totalEU

    sourceCscenario[j] <- sum(sourcescenario[1:ReportingCountryCount,j])
    percsourceCscenario[j] <- sourceCscenario[j]*100/totalEUscenario

    diffsourceC[j] <- sourceC[j] - sourceCscenario[j]
    diffpercsourceC[j] <- percsourceC[j] - percsourceCscenario[j]
    percdiffsourceC[j] <- diffsourceC[j]*100/sourceC[j]
}

# DEFINITION OF PRIORS

# Food-source-dependent factor (a): two-dimensional and uniform priors

for (c in 1:ReportingCountryCount)
{
    for (j in 1:FoodSourceCount)
    {
        a[c,j] ~ dunif(0,a_MaxRange[j])
    }
}

# Subtype-dependent factor (q): uni-dimensional and uniform priors

q[1] <- 1  # q for SE set to 1; all others will be estimated relative to this serovar
q[SeroVarCount] <- 1 # not used for estimating cases; value insignificant

for (i in 2:SeroVarCount-1)
{
    q[i] ~  dunif(0,q_MaxRange[1])
}


} # End of model
```

D5 – Variable definitions for Single Member State Model

```
[RequiredFields]
FoodImportExport=Year;Food_Source;Amount
Prevalence=Year;Genus;Species;SeroType;PhageType;GenoType;Food_Source;Units_Tested;Units_Positive
HumanCase=Year;Genus;Species;SeroType;PhageType;GenoType;Incidence;Travel_Yes;Travel_No;Travel_Unknown;Outbreak


[DimensionMinimums]
FoodSources=3
Subtypes=5
Years=2

[Distributions]
a=YearCount,FoodSourceCount
q=SeroVarCount
ptrav=YearCount,SeroVarCount

[Data]
Incidence=n[YearCount,SeroVarCount]
Travel_Yes=yt[YearCount,SeroVarCount]
Travel_No=nt[YearCount,SeroVarCount]
Travel_Unknown=pt[YearCount,SeroVarCount]
Outbreak=outb[YearCount,SeroVarCount]
ImportExport=m[YearCount,FoodSourceCount]
Prevalence=p[YearCount,FoodSourceCount,SeroVarCount]

[Distribution_a]
Grouping=FoodSourceCount
Default_Range=0,0.01
Visible:=True
FloatingPoint=True
UseMaxRangeArray=True
Transpose=False

[Distribution_q]
Grouping=
Default_Range=0,0.4
NA_Elements=1,SeroVarCount
Visible:=True
FloatingPoint=True
UseMaxRangeArray=True
Transpose=False

[Distribution_ptrav]
Grouping=YearCount
Default_Range=0,0.2
Visible:=True
FloatingPoint=True
UseMaxRangeArray=False
Transpose=False

[KernelDensityPlot]
a=Coda
q=Coda

[Winbugs_Defaults]
BurnInIterations=10000
RunIterations=30000
Chains=3

[Output_Variables]
```

source[YearCount,FoodSourceCount]=Number of estimated cases per source per year MS;MS;Baseline
travel[YearCount]=Number of estimated travel-related cases per year;MS;Baseline
unknown[YearCount]=Number of estimated cases with unknown source per year;MS;Baseline
total[YearCount]=Number of estimated total cases per year;MS;Baseline
percsource[YearCount,FoodSourceCount]=Percentage of estimated cases per source per year;MS;Baseline
lambdatji[YearCount,FoodSourceCount,SeroVarCount]=Number of estimated sporadic and domestic cases per serovar, source and year;MS;Baseline
lambdaexp[YearCount,SeroVarCount]=Number of estimated cases per serovar and year;MS;Baseline
a[yearcount]=Food-source related factor, one per source;MS;Baseline
q[SerovarCount=Serovar related factor, one per serovar. The q-value for the first serovar (S. Enteritidis) is by default set to 1 and the others q-values are estimated relative to this one.;MS;Baseline

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a contract between the European Food Safety Authority and the author(s), awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

D6 – WinBUGS code for Single Member State Model

```
Model
{
 for (t in 1:YearCount)
 {
  # Incl. others
  for (i in 1 : SeroVarCount)
  {
   o[t,i] <- n[t,i] - outb[t,i]
   ft[t,i] <- nt[t,i] + 1
   at[t,i] <- yt[t,i] + 1

   ptrav[t,i] ~ dbeta(at[t,i], ft[t,i])
   xtrav[t,i] <- ptrav[t,i]*pt[t,i]
   dt[t,i] <- xtrav[t,i] + yt[t,i]
   spdo[t,i] <- o[t,i] - dt[t,i]
  }
  # Excl. others
  for (i in 1 : SeroVarCount-1)
  {
   n[t,i] ~ dpois(lambdaexp[t,i])
   lambdaexp[t,i] <- lambda[t,i] + dt[t,i] + outb[t,i]
   lambda[t,i] <- sum(lambdatji[t,1:FoodSourceCount, i])
   uk[t,i] <- o[t,i] - lambdaexp[t,i]
   for (j in 1 : FoodSourceCount)
   {
    lambdatji[t,j,i] <- p[t,j,i]*m[t,j]*a[t,j]*q[i]
   }
  }



  for (j in 1 : FoodSourceCount)
  {
   source[t,j] <- sum(lambdatji[t,j,1:SeroVarCount-1])
   percsource[t,j] <- source[t,j] * 100/total[t]
  }
  travel[t]<- sum(dt[t,])
  unknown[t] <- spdo[t,SeroVarCount]  + sum(uk[t,1:SeroVarCount-1])
  outbreak[t] <- sum(outb[t,1:SeroVarCount])
  total[t]<- sum(source[t,1:FoodSourceCount]) + travel[t] + unknown[t] + outbreak[t]
 } # End for (t in 1:YearCount)


 for (t in 1:YearCount)
 {
   for (j in 1 : FoodSourceCount)
   {
    a[t,j] ~ dunif(0,a_MaxRange[j])
   }
 }

 q[1] <-1
 q[SeroVarCount] <-1
 for (i in 2:SeroVarCount-1)
 {
    q[i] ~ dunif(0,q_MaxRange[1])
 }
} #End Model
```