European Food Safety Authority

EFSA Journal 2015;13(1):3939

SCIENTIFIC OPINION

Scientific Opinion on the development of a risk ranking toolbox for the EFSA BIOHAZ Panel[1]

EFSA Panel on Biological Hazards[2,3]

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

Eight tools relevant to risk ranking of biological hazards in food were identified and assessed using two case studies. Differences in their performance were observed, related to the risk metrics, data requirements, ranking approach, model type, model variables and data integration. Quantitative stochastic models are the most reliable for risk ranking. However, this approach needs good characterisation of input parameters. The use of deterministic models that ignore variability may result in risk ranking errors. The ordinal scoring approaches in semi-quantitative models provide ranking with more errors than the deterministic approaches. FDA (Food and Drug Administration)-iRISK was identified as the most appropriate tool for risk ranking of microbiological hazards. The Burden of Communicable Diseases in Europe (BCoDE) toolkit can be used in combination with the outputs from FDA-iRISK or as a top-down tool to rank pathogens. Uncertainty needs to be addressed and communicated to decision makers and stakeholders as one of the outcomes of the risk ranking process. Uncertainty and variability can be represented by means of probability distributions. Techniques such as the NUSAP (numeral, unit, spread, assessment and pedigree) approach can also be used to prioritise factors for sensitivity and scenario analysis or stochastic modelling. Quantitative risk ranking models are preferred over semi-quantitative models. When data and time constraints do not allow quantitative risk ranking, semi-quantitative models could be used, but the limitations of these approaches linked to the selection and integration of the ordinal scores should be made explicit. Decision trees should be used only to show how decisions are made about classifying food–pathogen combinations into broad categories. BCoDE and FDA-iRISK, in combination with a network of available predictive microbiology tools, databases and information sources, can form a risk ranking toolbox and be applied based on a "fit for purpose" approach supporting timely and transparent risk ranking.

KEY WORDS

risk ranking, biological hazards, uncertainty, variability, quantitative tools

[1] On request from EFSA, Question No EFSA-Q-2013-00014, adopted on 4 December 2014.

[2] Panel members: Olivier Andreoletti, Dorte Lau Baggesen, Declan Bolton, Patrick Butaye, Paul Cook, Robert Davies, Pablo S. Fernández Escámez, John Griffin, Tine Hald, Arie Havelaar, Kostas Koutsoumanis, Roland Lindqvist, James McLauchlin, Truls Nesbakken, Miguel Prieto Maradona, Antonia Ricci, Giuseppe Ru, Moez Sanaa, Marion Simmons, John Sofos and John Threlfall. Correspondence: biohaz@efsa.europa.eu

[3] Acknowledgement: The Panel wishes to thank the members of the Working Group on the development of a risk ranking toolbox for the EFSA BIOHAZ Panel: Alessandro Cassini, Pablo S. Fernández Escámez, Arie Havelaar, Kostas Koutsoumanis, Roland Lindqvist and Moez Sanaa, the hearing expert: Fernando Pérez Rodríguez and EFSA staff: Alessandro Broglia, Winy Messens and Pablo Romero Barrios for the support provided to this scientific opinion.

Suggested citation: EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), 2015. Scientific Opinion on the development of a risk ranking toolbox for the EFSA BIOHAZ Panel. EFSA Journal 2015;13(1):3939, 131 pp. doi:10.2903/j.efsa.2015.3939

Available online: www.efsa.europa.eu/efsajournal

SUMMARY

The European Food Safety Authority (EFSA) asked the Panel on Biological Hazards (BIOHAZ) (i) to evaluate the performance and data requirements of the available risk ranking tools; (ii) to investigate methodologies for introducing uncertainty and variability in the risk ranking models; and (iii) to design and develop a risk ranking toolbox for the EFSA BIOHAZ Panel.

The BIOHAZ Panel identified eight tools relevant to risk ranking applications of biological hazards in food: decision trees; the United States Food and Drug Administration (US-FDA) risk ranking tool: the pathogen–produce pair attribution risk ranking tool (P^3ARRT); the EFSA food of non-animal origin risk ranking tool (EFoNAO-RRT); Risk Ranger; microHibro; swift quantitative microbiological risk assessment (sQMRA); FDA-iRISK; and the European Centre for Disease Prevention and Control (ECDC) Burden of Communicable Diseases in Europe (BCoDE) toolkit.

A detailed description of the tools, based on the conceptual risk ranking framework developed by the BIOHAZ Panel and their use in two risk ranking case studies, showed clear differences among them related to the risk metrics, the ranking approach, the model type, the model variables and data integration method. In addition, risk ranking tools have different data requirements, and empirical data requirements increase moving from qualitative to quantitative risk ranking approaches. When applied to the case studies of single pathogen–multiple foods (*Listeria monocytogenes* in ready-to-eat (RTE) foods) and multiple pathogens in a single food (leafy greens), the selection of the risk metric was found to significantly affect the risk ranking because the metrics measure different things, for example probability of illness versus public health burden (disability-adjusted life years (DALYs)). It should be noted that the performance of the risk ranking tools selected was evaluated from a statistical/theoretical perspective. Their implementation in practice may be constrained by limitations in data, time and resources.

Fully quantitative stochastic models are the most reliable for risk ranking. However, this approach needs a good characterisation of the input parameters. The evaluation of general approaches in risk ranking showed that the use of deterministic models that ignore variability may result in risk ranking errors, which may be greater for the food–pathogen combinations with the highest risk. When using semi-quantitative models with ordinal scoring, the food–pathogen combinations are classified into broad sets of categories with little discrimination. There are considerable differences in risk ranking compared with a quantitative stochastic model. The ordinal scoring approaches provide ranking with more errors than the deterministic approaches.

Among the quantitative tools that use a bottom-up approach for risk ranking, FDA-iRISK has been identified as the most appropriate for the needs of the EFSA BIOHAZ Panel. FDA-iRISK is a technically sound, quantitative tool providing meaningful risk metrics, allowing effective data management and scenario analysis. The evaluation of FDA-iRISK identified some limitations, including the omission of a maximum population density and the lack of uncertainty assessment. However, a new version of FDA-iRISK addressing most, if not all, of these issues will be available in the beginning of 2015. In addition, the BIOHAZ Panel concluded that BCoDE is a flexible, detailed and user-friendly DALY calculator that can be used in combination with the outputs from FDA-iRISK for a more effective calculation of DALYs or as a top-down tool based on epidemiological data to rank pathogens.

The BIOHAZ Panel evaluated methodologies to account for uncertainty in the risk assessment process. Uncertainty has been defined as "all types of limitations in knowledge, at the time it is collected". Uncertainty may arise from several factors in the risk ranking process, and includes technical (inexactness), methodological (unreliability), epistemological (ignorance) and societal (limited social robustness) aspects. Uncertainty in risk ranking needs to be carefully addressed and communicated to decision makers and stakeholders as one of the outcomes of the risk ranking process.

Uncertainty and variability can be represented in risk ranking by means of probability distributions, for example using two-dimensional Monte Carlo simulations. However, probabilistic representation is difficult when sufficient data are not available for statistical analysis. Expert elicitation procedures to incorporate diffuse information into the corresponding probability distributions may be adopted.

The NUSAP (numeral, unit, spread, assessment and pedigree) system aims to characterise and prioritise sources of uncertainty in a risk ranking model, and was used as an example of how to deal with uncertainty when using a risk ranking tool. NUSAP is a generic method that can be applied to all types of models and provides standardised scales for description of uncertainty in various dimensions. NUSAP uses expert judgement to evaluate the impact of uncertainty in individual model factors on the outcome of the assessment, leading to a prioritisation of factors for further work, for example sensitivity and scenario analysis, or stochastic modelling.

Quantitative risk ranking models respecting the rules of probability calculation and correctly describing the main biological phenomena that determine the risk are preferred over semi-quantitative models with ordinal scoring. When data and time constraints do not allow quantitative risk ranking, semi-quantitative models could be used. In this case, the limitations of these approaches linked to the selection and integration of the ordinal scores, as identified in this opinion, should be made explicit. Decision trees should be used only as a tool for showing how decisions about classifying pathogens–food combinations into broad categories are made (e.g. inclusion/exclusion; high/low). The BIOHAZ Panel concluded that BCoDE and FDA-iRISK, in combination with a network of available predictive microbiology tools, databases and information sources can form a risk ranking toolbox and be applied based on a "fit for purpose" approach supporting the timely and transparent development of risk ranking.

The BIOHAZ Panel recommended that the risk metrics used in risk ranking should have a meaningful biological or epidemiological interpretation and have to be agreed on by the risk managers before starting the risk ranking exercise. A framework encompassing uncertainty typology and evaluation (e.g. the NUSAP approach) should preferably be part of each risk ranking process to formalise discussions on uncertainty, considering practicality and feasibility aspects. In addition, a strategy should be developed to progressively adopt the proposed methods in future risk ranking opinions developed by the Panel.

TABLE OF CONTENTS

BACKGROUND

The setting of priorities plays a crucial role for the decision-making process in food safety management. In the face of finite resources, and a very large number of conflicting demands upon those resources, the establishment of priorities is a necessity.

Risk ranking is a technique that can be used to identify, and thereby prioritise, the most significant risks applying to a given situation. This methodology is also part of the overall EFSA Science strategy 2012–2016, and different complementary projects are running on this topic, involving different EFSA Units such as the Biological Hazards and Contaminants Unit (BIOCONTAM) and the Scientific Committee and Emerging Risks Unit.

The BIOHAZ Panel has already adopted scientific opinions where risk ranking was requested in the terms of reference, while the number of mandates that require a risk ranking exercise in the context of risk assessment is expected to increase in the future.

An opinion on the development of a conceptual risk ranking framework on biological hazards has been recently adopted by the BIOHAZ Panel (EFSA Panel on Biological Hazards (BIOHAZ), 2012b). In this opinion the risk ranking exercises relating to biological hazards undertaken in fourteen opinions that were produced by the BIOHAZ Panel were reviewed. It was concluded that there is no single and universally applicable standardised methodology for risk ranking. A conceptual risk ranking framework with nine separate stages was proposed to allow the adoption of the appropriate risk ranking methodology at each stage. Furthermore, nine risk ranking tools developed by other institutions worldwide were described, although none of these could be recommended as the single risk ranking tool for the BIOHAZ Panel.

In the adopted opinion it is also recommended that the development of a risk ranking toolbox based on the proposed framework should be undertaken, since such a toolbox would support the construction of consistent and transparent risk ranking models, and might assist the BIOHAZ Panel in the provision of timely answers to new mandates and food safety emergencies. The toolbox should be based on different modules that correspond to the nine stages of the framework with each module providing different option on risk metrics, ranking approaches, model types, variables and data integration methods. The above structure will allow the design and construction of risk ranking models targeted to the purpose of each mandate. At a first instance this toolbox will be of use for the BIOHAZ Panel, but the intention is that it will serve also for Members States, National food safety authorities and other food safety-related stakeholders.

In line with the above mentioned EFSA Science Strategy and as a complement to the BIOHAZ Panel's work, the Scientific Committee of EFSA launched a procurement call to perform a critical review of methodology and applications for risk ranking and benefit ranking for prioritisation of food and feed related issues, on the basis of the size of anticipated health impact. Although the latter call would not be limited to biological hazards, the results of this project is expected to provide additional insights on risk ranking and support the development of a risk ranking toolbox for the BIOHAZ Panel.

The overall objective of this self-task mandate in line with the European Commission priorities, is to capitalise on and advance the previous experience of the BIOHAZ Panel as well as the scientific and technical achievements for risk ranking through the development of a bespoke risk ranking toolbox. Developing such a tool and getting to the point of being able to apply it requires time and expertise, hence this proposal which will facilitate the provision of dedicated development time and resources for this important initiative.

TERMS OF REFERENCE

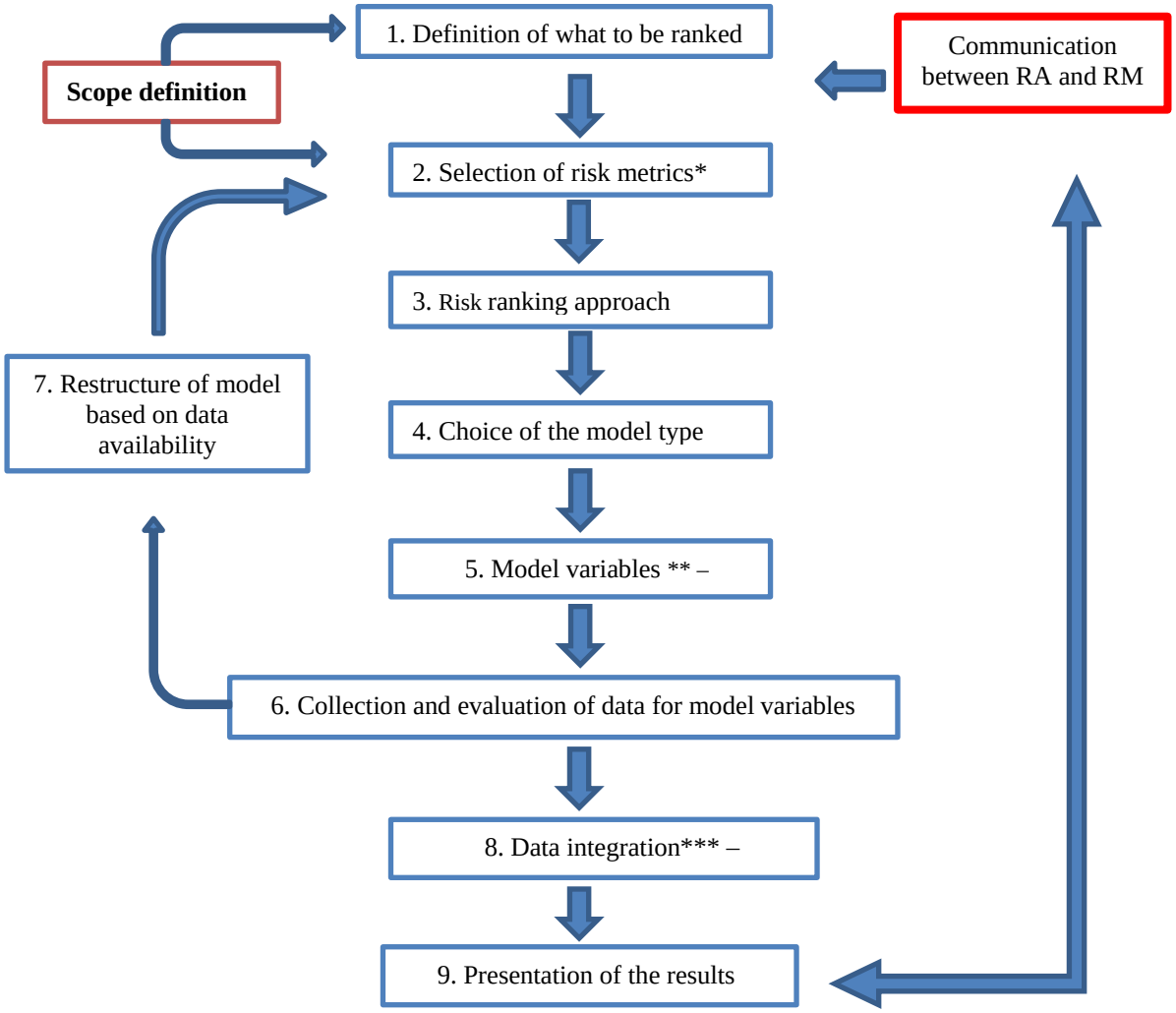
- To evaluate the performance and the data requirements of the available risk ranking tools.
- To investigate methodologies for introducing uncertainty and variability in the risk ranking models.
- To design and develop a risk ranking toolbox for the EFSA BIOHAZ Panel.

ASSESSMENT

1. Introduction

In the remit of the Panel on Biological Hazards (BIOHAZ), risk ranking is a coherent, comprehensive, transparent and evidence-based scientific process to prioritise and evaluate risks associated with biological hazards in foods. This aims to support decision makers in allocating resources to prevent and control health risks. Risk has been defined as "*a function of the probability of an adverse health effect and the severity of that effect, consequential to a hazard(s) in food*" (FAO/WHO, 2001).

In a previous opinion about the risk ranking framework, the BIOHAZ Panel proposed a conceptual risk ranking framework (Figure 1) comprising nine conceptual stages involved in risk ranking, from defining what is to be ranked to the presentation of the results of risk ranking (EFSA Panel on Biological Hazards (BIOHAZ), 2012b).



* "Risk metrics" is the expression of the risk (DALY – disability-adjusted life years, QALY – quality-adjusted life years, incidence, etc.).
** Model variables" are the indicators used for risk ranking (prevalence, epidemiological data).
*** "Data integration" is the combination of model inputs and formulas to produce model outputs.

RA: Risk assessment; RM: Risk management.

Figure 1: The proposed conceptual risk ranking framework for the BIOHAZ Panel

This framework shows that risk managers and risk assessors should be encouraged to liaise with each other regarding the aim of the risk ranking process and the communication of the results. It also provides the ability to adapt the appropriate risk ranking methodology by selecting different options at each stage. The appropriate option should be selected based on the aim of the risk ranking and available data. It was recommended that this conceptual risk ranking framework should be used in future risk ranking exercises in order to increase consistency and transparency. Furthermore, the proposed framework should represent the basis for the development of a risk ranking toolbox (i.e. a group of tools that could be used for risk ranking) since such a toolbox would support the construction of consistent and transparent risk ranking models.

In the previous opinion, nine risk ranking tools developed by institutions worldwide were identified and reviewed. They differed in their degree of complexity, level of quantification and approach to model construction.

The present opinion is a follow-up to the previous one and its scope is to carry out a comparative analysis of the performance of a selection of risk ranking tools on biological hazards and highlight their strengths and weaknesses. This exercise would allow the detection of the possible sources of uncertainty of different tools. In the timeframe of this opinion, the ultimate scope is to design a toolbox based on the risk ranking available tools with proper adjustments that cover the needs of the BIOHAZ Panel.

In order to address the terms of reference, the available risk ranking tools were identified and described in detail based on the conceptual risk ranking framework. The tools were further evaluated using two case studies: a single pathogen–multiple food setting (*Listeria monocytogenes* in ready-to-eat (RTE) foods) and multiple pathogens in a single food (Shiga toxin-producing *Escherichia coli* (STEC),[4] *Salmonella* spp., *L. monocytogenes*, *Campylobacter* spp., Norovirus, *Cryptosporidium* spp. and *Giardia* spp.) in leafy greens. The evaluation of the tools was based on their comparison with a fully quantitative stochastic risk ranking approach used in the reference model. In a next step, the general approaches in risk ranking, including a fully quantitative stochastic approach, a deterministic approach and a semi-quantitative approach with two scoring systems, were evaluated for various food–pathogen combinations using a common database. The incorporation of uncertainty and variability in risk ranking was explored. The use of the numeral, unit, spread, assessment and pedigree (NUSAP) approach for the identification of the important uncertainty sources was described and applied in a risk ranking case study. The methodologies of quantifying uncertainty in risk ranking models were also explored. The information gathered through these exercises was used to propose the tools to be included in a risk ranking toolbox. In addition, a prototype of a new risk ranking tool that covers the gaps of the current version of the available tools was developed. The use of these tools in a "fit for purpose" approach was presented. Finally, a supporting network of predictive microbiology tools, databases and information sources was presented as part of the risk ranking toolbox for the EFSA BIOHAZ Panel.

2. Description of selected risk ranking tools

From the tools that were evaluated in the previous opinion, some were considered to have too narrow a focus and therefore were not included in this assessment. In addition, some other tools have been recently developed, so in the end the following eight tools that can be used to rank the risk of microbiological hazards in foods were identified:

- decision trees;
- United States Food and Drug Administration (US-FDA) risk ranking tool: the pathogen–produce pair attribution risk ranking tool (P^3ARRT);
- EFSA food of non-animal origin risk ranking tool (EFoNAO-RRT);

[4] Shiga toxin-producing *Escherichia coli* (STEC) is also known as verotoxigenic *E. coli*, verocytotoxigenic *E. coli*, verotoxin producing *E. coli* and verocytotoxin-producing *Escherichia coli* (VTEC).

- Risk Ranger;
- microHibro;
- swift quantitative microbiological risk assessment (sQMRA);
- FDA-iRISK;
- European Centre for Disease Prevention and Control (ECDC) Burden of Communicable Diseases in Europe (BCoDE) toolkit.

Some of these tools can be classified as conforming to a "bottom-up" approach, that is, the agent is followed through the food chain to produce a prediction of risk to human health relative to other agents and/or foods. Other tools follow a "top-down" approach, where the level of risk associated with specific foods, hazards or their combinations is based on information gathered from epidemiological systems such as disease reporting and outbreak databases, while other tools combine both approaches (EFSA Panel on Biological Hazards (BIOHAZ), 2012b). A detailed description of the available risk ranking tools based on the conceptual risk ranking framework of EFSA is presented in the following paragraphs.

2.1. Decision trees

2.1.1. General description

Decision trees are simple tools that can be used for food safety risk assessment. The tool consists of a flow chart with alternative choices related to simple questions (typically with yes/no answers) allowing decisions to be taken, for example in a risk ranking. It is mainly based on qualitative inputs and it delivers a qualitative outcome. It is useful when the sources of information are qualitative or consist in poor quantitative data, providing very high versatility. The tool provides a qualitative indication of the risk associated with a food-borne hazard (categorised as, for example, high, medium or low). Owing to its simplicity, it can be adapted to the needs of the users.

2.1.2. Risk metrics

The metric associated with a decision tree in a "risk ranking", which provides a qualitative, categorised response of the relative risk associated with a hazard based (at least to a major extent) on qualitative information. The ranking establishes typically terms such as "high", "medium", "moderate", "low" and/or "negligible".

Decision trees allow a rapid comparison when there are many food-borne hazards to be considered and/or if there is a significant lack of quantitative information.

Decision trees are simple to use, although "expert opinion" is often needed to use them when scientific information is lacking.

2.1.3. Risk ranking approach

A decision tree can follow either a **"bottom-up"** (or forward) or a **"top-down"** (or backward) approach to public health risk ranking, adhering roughly to the standard microbial risk assessment paradigm in the former, or following a public health-based risk ranking that reflects the illness at the point of consumption in the latter.

2.1.4. Model type

Decision trees use a **qualitative** method for risk ranking. The tool requires the user to select from qualitative statements to provide a descriptive or categorical answer concerning factors that will affect the food safety risk to a specific or generic population.

Decision trees should take uncertainty and variability into account. Since in a qualitative approach there is no specific way in which uncertainty and variability in any input parameter are retained and reflected in the final risk estimate, the overall assessment can be evaluated in narrative, imprecise, terms such as "much", "little", etc., or scored according to the available evidence as in the case of evidence-based medicine. Another option is to include a number of scenarios that reflect the uncertainty and variability.

2.1.5. Model variables

The decision tree allows the most appropriate variables (major factors) that should be considered in decision-making to be established. The adoption of explicit variables allows participants to make a series of incremental judgements, which together can be combined to form an overall picture of the issue.

Typical input variables are (EFSA Panel on Biological Hazards (BIOHAZ), 2012a):

- hazard characterisation/identification;
- effect of process;
- effect of post-processing control system (Figure 2).

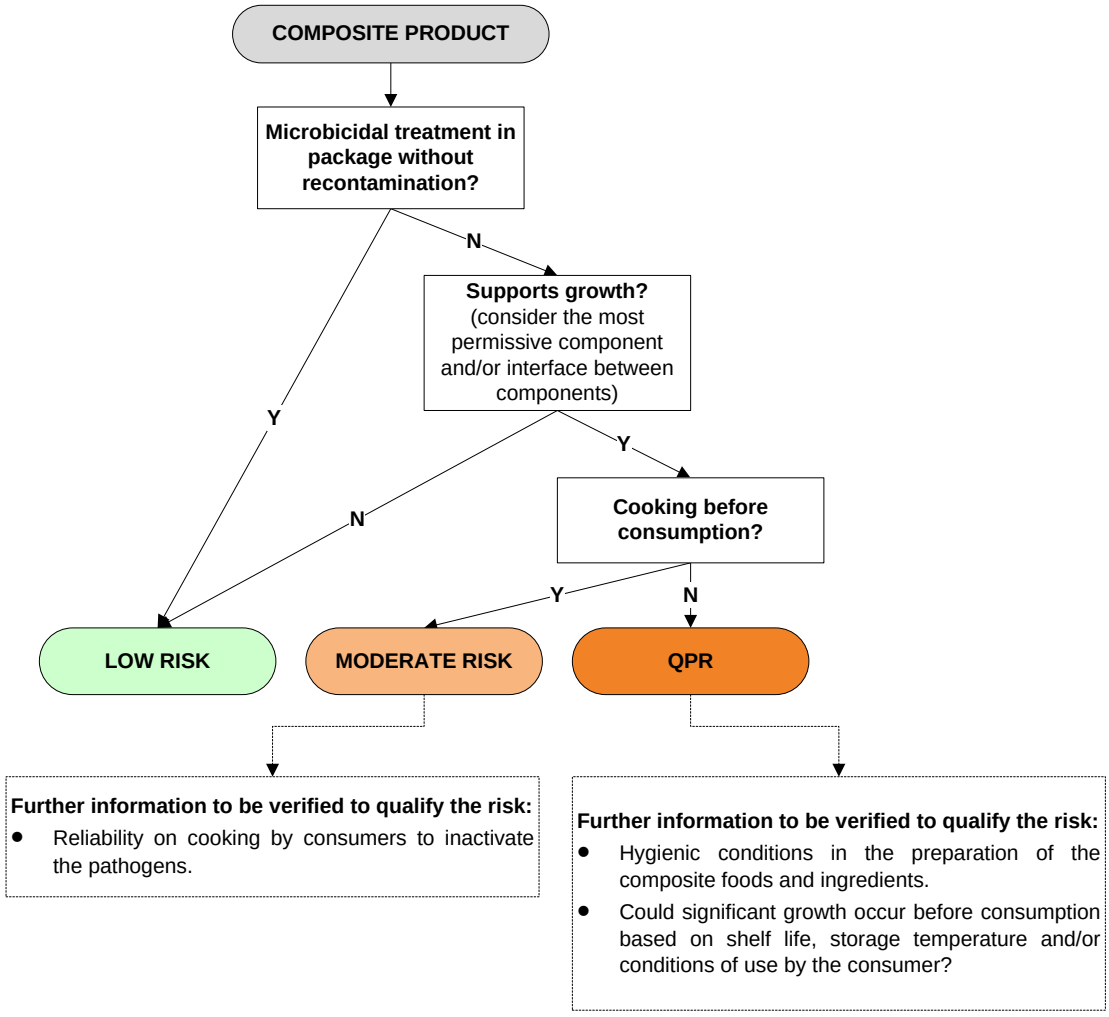
Variables can be selected according to the ranking needs. For most variables, scoring is based on categorical information as a response to simple questions which are relatively easy for the user to answer.

2.1.6. Data integration

As explained in the previous opinion, the data integration step combines information collected in the different stages of the risk ranking process (model inputs) to produce output results in the chosen risk metric (EFSA Panel on Biological Hazards (BIOHAZ), 2012b). Data integration in decision trees is based on a set of interconnected questions. Typically, little or no quantitative information is available.

2.1.7. Presentation of the results

The presentation of the results of a decision tree should be documented as fully as possible. This is particularly important as decision trees can present considerable variations and, in order to ensure transparency and reproducibility, the reasoning underlying the selection of different options must be explained in detail.



QPR: qualified presumption of risk.

Figure 2: Flow chart providing risk ranking of hazards which usually need to grow in food to cause illness (EFSA Panel on Biological Hazards (BIOHAZ), 2012a)

2.2. US-FDA risk ranking tool: the pathogen–produce pair attribution risk ranking tool or P^{3}ARRT

2.2.1. General description

The P^{3}ARRT is a semi-quantitative risk ranking software tool for prioritising, ranking and selecting pathogen–produce combinations (Anderson et al., 2011b). High-ranking combinations are prioritised for more rigorous risk assessment modelling efforts. Ranking is based on criteria related to the pathogen, human health and production/processing. A total of 11 data categories are used as input to estimate nine criteria—some data categories are combined into a single criterion. Data describe the strength of epidemiological association, severity of disease, pathogen characteristics that affect disease risk and commodity characteristics that affect pathogen prevalence, behaviour and likelihood of exposure by the consuming public. A total risk score is calculated for each of the selected pathogen–commodity pairs as the sum of the nine criteria scores multiplied by a corresponding criteria weighting. For the pathogen–commodity pairs included in the tool, a baseline ranking (default values) can be run and compared with ranking based on user-defined input. The user can define which criteria to include, the default bins for each data category, i.e. data limits for the four scoring bins (low, medium, high, very high) and the weight of each criterion. The tool is available for free download

(http://foodrisk.org/exclusives/rrt/) and is developed in Microsoft Access format, therefore this software is needed to run the tool.

2.2.2. Risk metrics

The P^{3}ARRT calculates one type of risk metric, the total risk score, which is the basis of the ranking list. The score is the sum of each criterion multiplied by a weighting factor. The weighting factor is included if the user considers one or more of the individual criteria more important than others. For each of the nine criteria, an ordinal number weight can be assigned. In the baseline criterion, a weight from 1 to 4 is used, but a weighting scheme from 1 to 100 or any range can be used (Anderson et al., 2011a). A criterion can be excluded from the ranking by entering a zero weight. Thus, the score can be evaluated by criteria category and, in addition, by changing weighting factors a sensitivity analysis is possible.

2.2.3. Risk ranking approach

The P^{3}ARRT follows a combined "bottom up" and "top-down" approach to public health risk ranking without explicitly referring to it in the description of the tool. The bottom-up part roughly adheres to the standard microbial risk assessment paradigm by inclusion of criteria related to prevalence of contamination, growth potential/shelf life; consumption; relative infectivity/infectious dose; susceptible population (reflecting risk groups and more severe consequences). The top-down approach/criteria reflect the public health burden by inclusion of criteria related to strength of epidemiological link, reflecting the extent of reported illness; epidemiological/disease, reflecting the "true" extent of illness; hospitalisation and death rates, reflecting the public health burden. The rationale for using selected criteria and approach for ranking in the tool is not explicitly explained other than no tool was available and the purpose was to design a transparent, data-driven, customizable, semi-quantitative, comparative risk assessment tool used to select priority pairs for further risk assessment efforts (Anderson et al., 2011a). Potential limitations of using a combined approach are not discussed in the tool. Limitations may include that some criteria (or sub-criteria) are related or correlated which may bias the scores used for ranking.

2.2.4. Model type

The P^{3}ARRT uses a **semi-quantitative** method for risk ranking, i.e. the quantitative data are divided into four categories, where each category is given a score, meaning that the final model outcome is presented on a semi-quantitative scale. The tool permits the user to refine the default bins for each of the four numerical scores associated with each criterion and to enter weights for each criterion. The Access software converts the quantitative data in the database tables into a category scored a value between 1 and 4 based on the defined bins. The criteria score is multiplied by the weighting factor and the total sum of the criteria included in the run is calculated by the software. The total sum score is the basis for ordering of the pathogen–commodity from high to low risk.

2.2.5. Model variables

The P^{3}ARRT includes 11 input variables (nine criteria) related to the risk or public health burden of pre-selected pathogen–commodity pairs. A total of 55 pathogen–commodity pairs are included in the tool. Pathogen–commodity pairs were selected by searching reports of outbreaks associated with fresh produce from the Annual Listing of Foodborne Disease Outbreaks compiled by the US Centers for Disease Control and Prevention (CDC) from 1996 to 2006, the Foodborne Outbreak Database sponsored by the Center for Science in the Public Interest, issues of Morbidity and Mortality Weekly Report (MMWR) from 1996 to 2008, and the peer-reviewed literature and publicly accessible databases. Only data from outbreaks of confirmed aetiology that occurred in the United States are included.

The input variables are (1) epidemiological link (number of outbreaks and total number of cases), (2) epidemiological multiplier (to account for unreported and undiagnosed cases), (3) hospitalisation (percentage of cases), (4) death (percentage of cases), (5) susceptible populations, (6) infectious

dose/relative infectivity, (7) contamination (Prevalence of contamination), (8) consumption (per cent of population consuming per day), and (9) shelf life and growth potential combined into one score.

2.2.6. Data integration

To generate the overall rank per pathogen–commodity pair, an algorithm that balances the score for each criterion with the weight of that criterion is used. The result is an overall numerical score for each pathogen–commodity pair that is produced by first multiplying each variable's score by its weight and then adding each of these nine values:

$$Rank\ score = \sum_{i=1}^{9} Score_i \times Weight_i$$

The algorithm is implemented in visual basic as a Microsoft Access database application, which can be run using Microsoft Access 2000, 2003 or 2007.

2.2.7. Presentation of the results

Tables in the database store raw data for each of the nine criteria and contain the parameter value at the commodity, pathogen or commodity–pathogen level. Other tables define the four bins (scored 1–4) for each of the nine criteria; these tables are linked to the raw data tables. The interface allows the user to reset the default bins for eight of the nine data categories (excluding the population susceptibility category and the growth potential category), as well as to determine appropriate weights for each of the nine data categories. When the user runs the application, ranking is performed based on user-specified inputs. A risk ranking summary report is generated that provides the list of pathogen–commodity pairs ordered by total score in descending order as well as a legend documenting the user inputs used to generate the list.

2.3. EFSA food of non-animal origin risk ranking tool (EFoNAO-RRT)—(adapted from the US-FDA risk ranking tool, P^{3}ARRT)

2.3.1. General description

EFoNAO-RRT was developed by the BIOHAZ Panel as a multi-criterion analysis model aimed at risk ranking combinations of food of non-animal origin commodities and specific pathogens (EFSA BIOHAZ Panel, 2013). It is a semi-quantitative tool that builds on the US-FDA P^{3}ARRT. Limited data availability, the use of broad risk categories and the possibility of applying qualitative or highly uncertain data were the stated reasons for developing an approach close to the P^{3}RRT model. The general modelling approach is a semi-quantitative risk ranking that takes into account variables such as the strength of association between the food commodity and the pathogen in question, the severity and extent of disease in humans and pathogen and commodity characteristics known to affect disease risk and/or probability of exposure. These variables are included in the model and used to define seven specific criteria that can be categorised as describing epidemiology and public health (criteria 1 to 3) as well as probability of exposure and risk (criteria 4 to 7). The model outputs of the tool are based on reported outbreaks associated with consumption of food of non-animal origin in the European Union (EU) Zoonoses Monitoring between 2007 and 2011. The model is implemented as a spreadsheet model in Microsoft Excel, which enables the user to modify data inputs and outputs.

2.3.2. Risk metrics

The EFoNAO-RRT calculates one type of risk metric, the total risk score. This score is the basis for ranking of pathogen–commodity pairs. For each criterion, available data were grouped into scoring categories, bins, which were defined and assigned a numerical, ordinal score. The total risk score is the sum of each criterion score multiplied by a weighting factor. Different weighting factors enable the user to consider one or more of the individual criteria more important than others. A criterion can be excluded from the ranking by entering a zero weight. Therefore, the score can be evaluated by criteria

category to investigate how robust the result is to different scenarios in terms of criteria included in the ranking or the weight put on the criteria.

2.3.3. Risk ranking approach

The EFoNAO-RRT follows a combined "bottom-up" and "top-down" approach to public health risk ranking. The bottom-up part roughly adheres to the standard microbial risk assessment paradigm by inclusion of the following criteria related to exposure and risk: prevalence of contamination, pathogen growth potential during shelf life; consumption; dose–response relationship. The top-down approach/criteria reflect the public health burden by inclusion of the following criteria: strength of epidemiological link, reflecting the extent of reported outbreaks; incidence of illness, reported cases corrected by a hazard-specific multiplier reflecting the "true" extent of illness; and burden of disease, reflecting the public health burden per 1 000 cases due to risk groups and more severe consequences.

2.3.4. Model type

The EFoNAO-RRT uses a semi-quantitative method for risk ranking. For each criterion, available quantitative data are grouped into defined scoring categories and assigned a numerical, ordinal score. The total sum of weighted criteria included in the run is calculated in the Excel spreadsheet. The total final risk score is the basis for ranking of all combinations. It is possible for the user to modify input data and how output is calculated in the spreadsheet model. For instance, definitions of scoring categories (bins) as well as the weights for each criterion can be modified. The total sum score is the basis for ordering of the pathogen–commodity from high to low risk.

2.3.5. Model variables

The EFoNAO-RRT includes input data for 10 variables used to categorise the seven criteria related to health consequences or risk of the pre-selected pathogen commodity pairs. A total of 32 pathogen–commodity pairs are included in the tool. The pathogen–commodity pairs were selected by identifying outbreaks associated with fresh produce from the reported food-borne outbreaks in EU Zoonoses Monitoring between 2007 and 2011. Only data from outbreaks classified as moderate to very strong (according to the number of cases) and that occurred in Europe are included. The criteria are: **(**1**) strength of associations between food and pathogen** (number of reported outbreaks and cases), **(**2**) incidence of illness** (notified number of cases and disease multiplier for under-reporting from EU *Salmonella* multiplier or multipliers, anchored to EU *Salmonella* (Scallan et al., 2011)), **(**3**) burden of disease** (DALY[5] per 1 000 cases based on data from the Netherlands (Havelaar et al., 2012)), **(4) dose–response relationship** (only three scoring levels), **(**5**) prevalence of contamination**, **(**6**) consumption** (percentage of consumers consuming, at least once, any specific food belonging to each EFoNAO category during the study period), and **(**7**) pathogen growth potential during shelf life** (combined score from growth potential and shelf life).

2.3.6. Data integration

The overall rank per pathogen–commodity pair incorporates all seven criteria scores and is estimated via an algorithm that balances the score for each criterion with the weight of that criterion. The result is an overall numerical score for each pathogen–commodity pair that is produced by first multiplying each variable's score by its weight and then adding each of these seven values:

$$Rank\ score = \sum_{i=1}^{7} Score_i \times Weight_i$$

The algorithm risk is implemented in Microsoft Excel.

[5] The DALY is a health gap measure that extends the concept of potential years of life lost as a result of premature death to include equivalent years of "healthy" life lost in states of less than full health or, in more general terms, disability. One DALY is one lost year of healthy life (World Health Organization definition). The DALY methodology has been described by Murray and Lopez (1994a, b, 1996) in the Global Burden of Disease (GBD) project.

2.3.7. Presentation of the results

For each combination of pathogen and commodity (rows), the score for each criterion as well as the total score are shown in columns in the spreadsheet. The separate columns allow the contribution of each criterion to be evaluated and ranking can be achieved by sorting based on the total score column.

2.4. Risk Ranger

2.4.1. General description

Risk Ranger is a simple tool for food safety risk assessment developed by the Australian Food Safety Centre (Ross and Sumner, 2002). The tool is in Excel spreadsheet format and embodies established principles of food safety risk assessment, i.e. the combination of probability of exposure to a food-borne hazard, the magnitude of hazard in a food when present and the probability and severity of outcomes that might arise from that level and frequency of exposure. The tool requires the user to select from qualitative statements and/or to provide quantitative data concerning factors that that will affect the food safety risk to a specific population, arising from a specific food product and specific hazard, during the steps from harvest to consumption. The spreadsheet converts the qualitative inputs into numerical values and combines them with the quantitative inputs in a series of mathematical and logical steps using standard spreadsheet functions. These calculations are used to generate indices of the public health risk.

2.4.2. Risk metrics

Three types of risk metrics are calculated in Risk Ranger. The first is the "**probability of illness per consumer per day**", calculated as $P_{inf} \times P_{exp}$, where P_{inf} is the probability of a disease-causing dose being present in a portion of the product of interest and P_{exp} is the probability of exposure to the product per person per day. This metric is not strictly a measure of risk, because it does not include the severity of the illness resulting from exposure to the hazard. The second metric is the "**total predicted illnesses/annum in population of interest**", which does not differentiate severity either, but provides another measure that might be more readily understood than risk per day. The third metric is the "**risk ranking**", which provides a more user-friendly and robust index of relative risk and is calculated based on the "comparative risk" estimate. The "comparative risk" in the population of interest is a measure of relative risk which includes the severity of the illness and is independent of the size of the population, but does consider the proportion of the population consuming. A "comparative risk" of 1 represents the situation in which every person in the population consumes the product of interest daily, and that each portion of the product contains a lethal dose of the hazard. The "risk ranking" value is scaled logarithmically between 0 and 100, where 0 represents no risk and 100 represents the opposite extreme, where every member of the population eats a meal that contains a lethal dose of the hazard every day. The "risk ranking" scale is set based on a probability of mild food-borne illness of less than or equal to one case per 10 billion people (greater than current global population) per 100 years as a negligible risk. The "comparative risk" estimate that corresponds to this value is $2.75 \times 10 - 17$ and the "risk ranking" corresponding to this level is equated to zero. Analogously, the upper limit of "risk ranking" at 100 corresponds to a "comparative risk" of 1.

2.4.3. Risk ranking approach

The Risk Ranger follows a "**bottom-up**" (or forward) approach to public health risk ranking, corresponding roughly to the standard microbial risk assessment paradigm. The risk ranking is based on factors that affect the food safety risk to a specific population, arising from a specific food product and specific hazard, during the steps from harvest to consumption.

2.4.4. Model type

Risk Ranger uses a **semi-quantitative** method for risk ranking. The tool requires the user to select from qualitative statements and/or to provide quantitative data concerning factors that that will affect the food safety risk to a specific population, arising from a specific food product and specific hazard,

during the steps from harvest to consumption. The spreadsheet converts the qualitative inputs into numerical values and combines them with the quantitative inputs in a series of mathematical and logical steps using standard spreadsheet functions. These calculations are used to generate indices of the public health risk.

2.4.5. Model variables

The Risk Ranger model includes 11 input variables related to the severity of the hazard, the likelihood of a disease causing dose of the hazard being present in a meal and probability of exposure to the hazard in a defined period of time. These variables are: (1) **hazard severity**, (2) **susceptibility of the consumer**, (3) **frequency of consumption**, (4) **proportion of population consuming**, (5) **size of population of interest**, (6) **proportion of product contaminated**, (7) **effect of process**, (8) **potential for recontamination after processing**, (9) **effect of post-processing control system**, (10) **increase from level at processing required to reach an infectious or toxic dose for the average consumer**, and (11) **effect of preparation for meal**. For most variables, scoring is based on ordinal weighting factors translated to simple questions which are relatively easy for the user to answer.

2.4.6. Data integration

The logic and equations leading to the risk estimates are detailed below.

The "**probability of illness per consumer per day**" is calculated as $P_{inf} \times P_{exp}$, where P_{inf} is the probability of a disease-causing dose being present in a portion of the product of interest and P_{exp} is the probability of exposure to the product per person per day.

P_{inf} is defined as whichever is the larger value of the product of the values of the following variables (V):

> proportion of product contaminated (V6) × effect of process on the probability of contamination (V7) × effect of post-processing handling/storage (V9) × increase in the initial level of the factor required to reach an infectious dose (V10) × effect of preparation prior to eating (V11)

or, in the case of a process resulting in the elimination of the hazard:

> proportion of product re-contaminated (V8) × effect of post-processing handling/storage (V9) × increase in the initial level of the factor required to reach an infectious dose (V10) × effect of preparation prior to eating (V11)

P_{exp} is given by the product:

> frequency of consumption (V3) × proportion of the population that consumes the product (4)

The "total predicted illnesses/annum in population of interest" is calculated as:

> 365 (i.e. days per year) × "probability of illness per consumer per day" (as described above) × fraction of population considered in at risk class (V2) × total population (V5)

The "risk ranking" is calculated based on the "comparative risk" estimate:

> probability of illness per day per consumer of interest (as described above) × hazard severity (V1) × proportion of population consuming (V4) × proportion of total population in population of interest (V2)

2.4.7. Presentation of the results

The results of Risk Ranger tool are the values of the three risk metrics described above which are presented in a simple Excel spreadsheet together with the selected values of the variables. The user has

to calculate the above risk metrics for each pair of food–pathogen separately and compare them manually. There is no graphical representation of the results.

2.5. microHibro

2.5.1. General description

microHibro is a quantitative model based on prevalence and concentration data for pathogens at the starting point of the risk assessment and then using cross-contamination, growth, survival, intervention rates and dose–response as key variables that would affect the outcome of the model.

microHibro is the result of a national and regional project focused on RTE products, which is being expanding to new food categories. The main purpose is to offer an easy-to-use tool to end-users, risk managers, food business operators and risk assessors. The application has been developed as a web-based tool, considering as important features, flexibility, updatability and usability, underpinned in a solid and validated mathematical structure. The software is developed by Hibro research group (University of Cordoba, Spain) in collaboration with Optimum Quality, a software developing spin-off company (Technological Park of Cordoba, Spain). Hibro is in charge of administering, promoting and improving microHibro software from both technical and applicative sides.

The mathematical structure of the exposure assessment model was translated into a user-friendly web-based tool which is released as a beta version in English and Spanish (microHibro 2.0 Beta, www.microhibro.com). The software is licensed as General Public Licence (GPL) or equivalent, with open access. The application incorporates a module for growth predictions in different vegetable matrices and microorganisms as well as a module which allows the user to design and simulate exposure models to estimate the final concentration at the moment of consumption. It allows models selected by the user to be introduced into the application.

microHibro is a stochastic modelling tool whose risk modelling module can be used for risk assessment, incorporating deterministic or quantitative values for initial concentration, growth, inactivation, recontamination and dose–response. Information about the variables can be included as either deterministic or stochastic data. The flexibility of the tool would allow the addition of further components. It provides an estimation of the risk and the probability of disease. Finally, the sensitivity analysis tools can be then applied to assess how variables and factors can impact the number of cases, i.e. public health.

2.5.2. Risk metrics

By inserting variables as point-estimate values or distributions, microHibro calculates outputs as frequency distribution of microbial growth and of probability of illness. The risk metric used is the "probability of illness".

2.5.3. Risk ranking approach

microHibro follows a bottom-up (or forward) approach considering the steps of initial microbial concentration (including prevalence), growth, transfer, reduction and dose–response.

2.5.4. Model type

microHibro is a quantitative model for simulating growth of microorganisms in food matrices and estimating the probability of illness.

2.5.5. Model variables (inputs)

The model allows the carrying out of a probabilistic exposure assessment based on an object-oriented system with four model variables, i.e. growth, inactivation, transfer and dose–response, that can be defined by using either point-estimate or probability distributions of mass, temperature, pH, time, etc. The types of distributions include continuous (normal, exponential, uniform and triangular) and

discrete (binomial and Poisson) distributions. In the case of continuous distributions, the concentration unit is $\log_{10}$ colony-forming units (CFU); however, the discrete distributions, because of their discrete nature, are defined by arithmetic units, i.e. CFU.

Distributions are defined by giving values to the parameters of the selected distribution. For example, in the case of normal distribution, the parameters to be defined are the standard deviation and mean. These are the input elements that can be selected:

- Element 1: initial concentration, mass and prevalence. The initial concentration and prevalence can be implemented as distributions selected from a list or as fixed values, whereas mass can be included only as a fixed value.
- Element 2: growth. The user can choose: (1) a selection of published models available, (2) to include additional models or (3) to introduce a distribution among the ones available in the tool or to include a fixed value. The mass can also be included.
- Element 3: microbial transfer. Information about cross-contamination can be implemented. In order to do so, either distributions (from a list of continuous and discrete ones) or fixed values of the percentage of transfer of microorganisms and microbial concentration can be selected. The mass and probability of occurrence can also be included as fixed values.
- Element 4: reduction in the concentration of microorganisms. Factors meaning a decrease in microbial concentration can also be considered. The user can choose: (1) a selection of published models available, (2) to include additional models or (3) to introduce a distribution among the ones available in the tool or to include a fixed value. The mass can also be included.
- Element 5: dose–response models. There are dose–response models available in the tool or the user can implement new models.

2.5.6. Data integration

microHibro allows a probabilistic exposure assessment to be carried out based on object-oriented system with three basic types of predictive model: growth, inactivation and cross-contamination. The simulation method used in the application is based on the Monte Carlo method, which allows the generation of random values from defined probability distributions. To this end, the inversion method for generating random numbers was applied (Robert and Casella, 2004). A detailed user manual is available in the website (www.microhibro.com).

2.5.7. Presentation of the results

The application includes a basic tool for sensitivity analysis that allows: (1) the comparison of the inputs and outputs data graphically and (2) the simulation different scenarios of the designed risk model.

With the first option, the simulated values for input variables, such as temperature, pH, etc., are plotted versus concentration and prevalence outputs derived from the simulated risk model in a scatter plot.

With the scenario analysis, the effect of specific input variables (temperature, pH, etc.) on the final concentration of microorganism (i.e. output) is quantitatively assessed. The information obtained by the sensitivity analysis may be used to identify critical control points and risk thresholds in the analysed input variable. The application performs a set of simulations, each using one of the defined values, which is fixed during the whole simulation while the remaining variables are allowed to vary. The application returns graphs representing the changes in the main statistics of the final output with respect to the values specified for the analysed input variable.

2.5.8. User interface

This tool is based on an object-oriented system. The types of objects are in the bar of design, at the bottom of the application and represent the previously mentioned basic models (growth, inactivation, cross-contamination). The user can design a processing line or specific food chain by dragging these objects to the design space in the central part of the application. The risk model is represented in the design space as a flow chart, where each basic model stands linearly behind another, according to how they were initially placed. It is a very user-friendly, easy-to-use interface.

2.6. Swift quantitative microbiological risk assessment (sQMRA) tool

2.6.1. General description

sQMRA is a tool for food safety risk assessment developed by the Dutch National Institute for Public Health and the Environment. The tool is in Excel spreadsheet format and is based on general principles of food safety risk assessment, providing a standardised environment for full quantitative risk assessment. The model covers the food chain from retail to preparation and consumption and carries on to infection and illness. The first version (Evers and Chardon, 2010) is deterministic and calculates the probability of illness for a pathogen–product combination by estimating the exposure to a food-borne hazard for a number of categories which are input for a dose–response relationship. Recognising the limitations of simplified QMRA models, the tool is designed primarily for comparative risk assessment, regarding both the final risk estimates and the intermediate results. The second version of the tool (sQMRAv2) is stochastic (considering variability but not uncertainty) and is implemented using the @RISK add-in to Excel (Evers and Chardon, 2012, 2013). This second version was used in this assessment, as it included many improvements, such as growth or inactivation during storage by the consumer, an extended cooking module, a choice of two dose–response models, extended results presentation and reference and user-defined comparison datasets. The second version also provides estimates for severity of illness, using DALYs and cost-of-illness. The tool requires the user to enter data on prevalence and concentration of pathogens at retail, food consumption, effects of storage, cooking and cross-contamination in the kitchen, a dose–response relationship and, in version 2, on disease burden and cost of illness. The spreadsheet then converts the inputs into risk estimates using established algorithms.

2.6.2. Risk metrics

Several types of risk metrics are calculated by the sQMRA tool. The description below focuses on the second version of the tool. The output sheet of the tool provides summary information, whereas the model sheet provides extended intermediate calculation results:

- Summary information on the scope of the risk assessment and input parameters.
- Attribution of exposure (probability of exposure and total exposure) and illness estimates over different categories of storage by the consumer (room temperature, fridge or freezer).
- Attribution of exposure (probability of exposure and total exposure) and illness estimates over different pathways in the kitchen (survival of heating well-cooked, undercooked or raw; cross-contamination).
- Contamination level (prevalence and number) at portion and population level in several steps of the food chain.
- Infection, illness, disease burden and cost-of-illness at portion and population level.
- Variability of contamination and effect estimates at portion level.
- Statistical uncertainty in food consumption and retail data is included for illustrative purposes, but not used in the model calculations.

2.6.3. Risk ranking approach

The sQMRA tool follows a "**bottom-up**" (or forward) approach to public health risk based on the standard microbial risk assessment paradigm, but restricted to the retail-to-consumption part of the food chain. Key outputs for risk ranking are:

- contamination level (prevalence and number) at portion and population level in several steps of the food chain, and compared with a chosen reference model;
- infection, illness, disease burden and cost-of-illness at portion and population level, and compared with a chosen reference model.

2.6.4. Model type

sQMRAv2 uses a **quantitative, stochastic** method for risk ranking. The tool requires the user to provide quantitative data concerning factors that will affect the food safety risk for consumers, arising from a specific food product and specific hazard, during the steps from retail to consumption.

2.6.5. Model variables

The sQMRA tool model includes 14 categories of input variables. These categories are:

- portions consumed
- pathogen prevalence in retail
- portion size
- pathogen concentration
- storage conditions
- growth and inactivation characteristics of pathogen
- cross-contamination parameters
- preparation categories
- probability of survival during preparation
- endpoint dose–response model
- dose–response parameters
- probability of illness given infection
- DALY per case
- cost-of-illness per case.

For all variables, variability distributions are optional. The user can also enter deterministic information by leaving the cells for standard deviations (and other variability statistics) blank.

2.6.6. Data integration

The model equations are fully described and follow standard QMRA methodology.

2.6.7. Presentation of the results

The results of sQMRA tool are the values of the risk metrics described above, which are presented in the RESULTS sheet. Several summary graphs and tables are available. The built-in graphical presentations focus on comparison of the risk in relation to different storage conditions and preparation methods for the food product. For risk ranking, seven pre-defined reference datasets are available in the tool, with the CARMA model for *Campylobacter* on broiler meat (Nauta et al., 2007)

offered as default. The user can also enter additional scenarios. Once a user-defined reference pathogen–product combination is available in the tool, it can be selected with a drop-down list in the RESULTS sheet for 1:1 comparisons with the model scenario.

2.7. FDA-iRISK

2.7.1. General description

The FDA-iRISK is a comparative risk assessment system for evaluating and ranking food–pathogen pairs developed by the Food and Drug Administration (FDA) through partnership and collaboration with experts within and outside US government organisations. It is designed to estimate risks associated with both microbial and chemical hazards (Chen et al., 2013).

The FDA-iRISK is a web-based quantitative risk assessment system that enables users to assess, compare and rank the risks linked to multiple food–pathogen pairs: a relative rapid quantitative risk assessment. It is a modelling tool that integrates data on the hazard, data on the food supply system (from primary production, through manufacturing and processing, to retail distribution) and data on dose–response and health effects, using the built-in mathematical logic/equations and Monte Carlo simulations. It enables also evaluation of impact of interventions applied all over the food supply.

The web-based user interface enables users to define the food and the hazard of interest, edit inputs, update references and assumptions, and store, view and share data, information and risk scenarios. The version used in this assessment was FDA-iRISK 1.0 (hereafter referred to simply as FDA-iRISK).

2.7.2. Risk metrics

To enable the comparisons of risks posed by different food–pathogen pairs, iRISK is using DALY as a common metric. DALY is an indicator of the time lived with a disability and the time lost because of premature mortality. A DALYs per case value is used as a measure of the averaged burden of disease per case of illness associated with each hazard, taking into account the relative frequency of each potential health outcome. The final output of FDA-iRISK, the annual DALYs, is obtained by multiplying the DALYs per case by the annual expected number of cases for a food–pathogen pair under evaluation.

2.7.3. Risk ranking approach

The FDA-iRISK follows a “bottom-up” (or forward) approach to public health risk ranking adhering to the standard microbial risk assessment paradigm. Risk ranking is based on factors that affect the food safety risk to a specific population, arising from a specific food product and specific hazard, during the steps from primary production to consumption.

2.7.4. Model type

The FDA-iRISK uses a quantitative method for risk ranking. The tool requires the user to specify hazards, foods and populations of interest and inputs data related to the exposure assessment and hazard characterisation components as defined in CODEX risk assessment standard.

The FDA-iRISK provides a risk assessment model framework and templates, and the user chooses the suitable template for her/his risk scenario and provides evidence including the possibility for documenting the rationale behind the selection of the evidence. A risk scenario is defined by seven elements, described below.

2.7.5. Model variables (inputs)

The FDA-iRISK tool includes seven input elements:

- Element 1: foods. The definition of food and its description will affect the process model.

- Element 2: hazards. The type of hazard will affect process model options and dose–response options provided within FDA-iRISK for the hazard.
- Element 3: population groups. The choice of the population group is associated with the choice of dose–response model (e.g. two dose–response models for *L. monocytogenes*, one for high-risk population and another for low-risk population), specific patterns of health effects (e.g. pregnant women for abortion) and the consumption patterns (e.g. specific diet per age group).
- Element 4: process models. The process model describes the impact of the different process stages (primary production, food processing, food handling, etc.) on the concentration and prevalence of the hazard in the considered food. The outputs of the process model are the probability distribution of the concentration of the hazard in a food serving and the prevalence of contaminated servings. Data required include the initial prevalence, distribution of the hazard concentration and the unit mass, data related to process stages from farm to table of the food supply chain up to the point of consumption. The number of stages depends on the food definition, hazard characteristics and the scope of the risk assessment. For example, the initial prevalence and concentration could be at retail level or at the primary production level. Hence, the process model is designed as a series of process stages, events or steps along the farm-to-fork continuum. At each process stage, the user provides the expected impact of the considered stage on the prevalence and concentration of the hazards and on the unit size of the food. The effect, such as increase/decrease of the prevalence, increase/decrease of the hazard concentration in food, can be expressed as a fixed value or as a probability distribution. The process types and their data inputs are described in Appendix A. The template proposes nine process types: (1) increase by growth; (2) increase by addition (as cross-contamination from the processing environment); (3) decrease; (4) pooling; (5) partitioning; (6) evaporation or dilution; (7) partial redistribution that models partial cross-contamination among food units. The total hazards load remains constant; (8) total redistribution: the total hazards load is redistributed to all food units; (9) no change.
- Element 5: consumption models. The consumption models are defined in relation to the specific population groups. For microbial hazards, the required inputs are the serving size (fixed value or distribution) per each food eating occasion and the number of eating occasions per year. For chemicals, the distribution of the average amount of the food eaten daily over a period of time or a lifetime and the number of consumers are required.
- Element 6: dose–response models. The dose–response models are defined in relation to the specific population groups.
- Element 7: health outcomes.

2.7.6. Data integration

The equations leading to the risk estimates are not detailed in FDA-iRISK technical documentation. In relation to the integration of the different elements of the risk assessment, especially for process types, it is said that the implemented program uses previously published mathematical equations by the International Life Sciences Institute (2010) and Nauta (2005, 2008) without mentioning them. At this step of FDA-iRISK reviewing, it is not possible to describe precisely how FDA-iRISK integrates all the different process types (see Appendix A). However, further to the information existing in the technical documentation, a paper by Chen et al. (2013), and checking the model outputs, the following integration procedure can be proposed, when the final concentration and prevalence are assessed (see graphical representation in Appendix A):

1. A dose distribution per serving is generated taking into account the variability of the hazard concentration at the time of consumption (among contaminated servings) and the variability of serving size (among consumers from the same population group).

2. A risk-per-serving distribution is then derived using the population group-specific dose–response model.
3. The arithmetic mean of risk per contaminated serving is calculated from the risk per serving distribution (step 2). The mean risk of illness per serving is then calculated by multiplying the mean of risk per contaminated serving by the prevalence of contaminated units at the time of consumption.
4. The expected annual number of cases of illness is then calculated by multiplying the mean risk of illness per serving by the total annual number of eating occasions for the considered population group.
5. Finally, the annual DALYs are derived by multiplying the expected annual number of cases of illness by the DALYs per case of illness.

Steps 1 to 5 of inputs integration are conducted per population group. The final output is the sum of DALYs obtained for each population group.

2.7.7. Presentation of the results

The outputs of the FDA-iRISK annual cases and annual DALYs, as well as the inputs, are summarised in tables provided in portable document format. Before creating the final report a filtering system using the description of the different scenarios is proposed to enable the different possibilities of ranking (comparisons).

2.7.8. User interface

FDA-iRISK is a web-based free software. It uses a tabbed interface to provide access to its functionality. When clicking on a link or tab, FDA-iRISK opens the requested page in the current window. Only one window is used at a time. Before using FDA-iRISK it is recommended to follow the quick start tutorial (downloadable from the FDA-iRISK website). This tutorial is very useful to understand the logic of the nested folders and the definition of the scenarios. The organisation of the different folders is intuitive for users with a minimum background on quantitative risk assessment.

2.8. ECDC Burden of Communicable Diseases in Europe (BCoDE) toolkit

2.8.1. General description

The BCoDE is a project led and funded by the European Centre for Disease Prevention and Control (ECDC) and by a European consortium with the purpose of estimating the impact of communicable diseases (CDs) in the EU and the European Economic Area (EEA) Member States (MS). The project has four main objectives: (1) to promote evidence-based methods in epidemiology and decision-making; (2) to introduce tools for planning and prioritisation; (3) to identify gaps in surveillance data availability and quality; (4) to provide tools for communicating complex information to decision makers.

The BCoDE project has developed a stand-alone software providing a user-friendly interface, the BCoDE toolkit. The toolkit allows the calculation of the burden, expressed in DALYs of 32 CDs of interest to ECDC. This is made available to MS national experts to allow the estimation of national burden of CDs. The aim of the software is to assist MS in applying the proposed BCoDE evidence-based approach for estimation of the burden of CDs, and to facilitate communication between data generators and users through multiple visualisation options, ultimately fostering its value in health policy formulation. The application is written in C++ using Qt C++ toolkit, version 4.8.4. All computations are implemented in C++ and the interface is HTML with JavaScript.

Each selected disease generates a model visible as a graphical outcome tree. By default, users input country-specific notified data (optional The European Surveillance System (TESSy) data source) and age- and gender-specific multiplication factors, adjusting for under-estimation. The BCoDE toolkit

requires input of cases for 19 age groups (20 in case the congenital form is relevant, e.g. listeriosis) and gender split (overall 38/40 inputs). If necessary, the user is also allowed to edit population data as well as parameters of the outcome tree. The software will estimate the burden based on disease models describing the natural history of the disease, ensuring sequelae are considered. Calculations are based on Markov models and the number of iterations is chosen by the user. The output phase displays disease-specific results, impact of acute illness versus sequelae, gender- and age-specific DALYs and DALYs per 100 000 persons with uncertainty intervals, years of life lost as a result of premature mortality (YLLs), years lived with disability (YLDs) and DALYs per case.

Aggregated results enabling a comparative assessment of the impact of CDs are displayed as bubble charts (DALYs/100 000) plotted against mortality, incidence and DALYs/case. Interactive tables and bar charts ranking diseases and uncertainty can be produced and exported.

2.8.2. Risk metrics

The main risk metric produced by the BCoDE toolkit is the burden of disease expressed in DALYs. Other related metrics include DALYs per 100 000, DALYs per case, YLD and YLL. These are available for each disease, age group, sex and outcomes of a given disease (the latter also includes data on incidence and mortality). Uncertainty intervals (lower bound 2.5th percentile and upper bound 97.5th percentile), median and mean are displayed next to all the above-mentioned outputs.

2.8.3. Risk ranking approach

The methodology underlying the BCoDE toolkit is incidence and pathogen based, which basically entails a top-down approach. Risk ranking of infectious diseases involves listing these according to their impact on population health. For ranking of food-borne pathogens, additional data on attribution of the total disease incidence in a population to exposure by food in contrast to other pathways are necessary. If the risk assessment question is at a more detailed level (e.g. ranking the hazards in leafy greens), even more detailed attribution data are necessary. Such data are currently difficult to obtain for all food-borne hazards in the EU.

Of interest for the purpose of the opinion, the BCoDE toolkit allows flexibility relating to population data (main denominator) and incidence (main input). For example, numerous models of the same disease can be created (and at will, changing the populations of the different models), allowing scenario analysis expressing the risk of a given food or different foods or categories of food, according to their risk of infecting humans in the same or different populations. This allows the tool to be used in a bottom-up approach, by using incidence estimates from the outputs of a quantitative risk assessment tool (e.g. FDA-iRISK, sQMRA) as inputs for the BCoDE tool. The former tools can provide incidence for a limited number of age groups (generally < 1 year of age, > 60 years of age and in between), whereas the BCoDE toolkit requires input of cases for 19 age groups (20 in case the congenital form is relevant, e.g. listeriosis) and male/female (overall 38/40 inputs). The user can redistribute the FDA-iRISK and sQMRA outputs according to the observed age/sex distribution in the EU/EEA as reported to the ECDC TESSy database. This combination of tools would create a very flexible, powerful and detailed approach to rank risks of pathogens in food at any desired level of detail with regard to food–pathogen combinations.

2.8.4. Model type

The BCoDE toolkit uses a quantitative method for risk ranking. Both the default input and the optional changes to the models are quantitative: number of cases, population under study, life expectancy and disease model parameters are numerical and based on the best available evidence. The output is also fully numerical, with uncertainty expressed in confidence intervals.

2.8.5. Model variables

Overall, seven sets of input variables are editable. Default variables are age and sex specific (1) number of cases and (2) multiplication factors adjusting under-estimation. All other embedded

variables are editable per gender and age group: (3) population, (4) life expectancy and outcome tree parameters (5) transitional probabilities, (6) disability weights and (7) duration of outcomes.

2.8.6. Data integration

Each disease is represented by a model including the related health outcomes and its burden is calculated and expressed in DALYs. These are not age-weighted and no time discounting is applied.

2.8.7. Presentation of the results

Results are printable and exportable in portable document format and in Excel.

Results are presented in two tabs:

- **Detailed results**: a specific page presenting all results relative to each specific disease. The user can browse diseases across a scroll-down menu. For each disease the following is presented:
  - Results table. Columns include the following: cases, incidence, YLD, YLL and DALYs per year, DALYs per case and DALYs per 100 000. Rows include: total infected, total acute, each sequela included in the outcome tree (also cases and deaths) and total sequelae.
  - Two coloured results charts. A bar chart including total DALYs, DALYs due to acute disease and DALYs due to sequelae, all split between YLL and YLD. A pie chart summarising the contribution of each sequela and acute to the total burden of the disease.
  - Results details. Table with age group- and sex-specific results, including 2.5^{th} and 97.5^{th} percentiles for DALYs, DALYs per 100 000 and DALYs per case. Bar chart with age group- and sex-specific results and percentile intervals.
- **Aggregated results**: results for all diseases are summarised; comparability and ranking is allowed.
  - Mortality/incidence comparison. Static bubble chart: each disease is represented by a bubble of a different colour. The diameter of the bubble represents DALYs per 100 000. Each bubble is plotted against incidence per 100 000 (x-axis) and mortality per 100 000 (y-axis). An interactive legend is available on the right-hand side: if a disease is unselected, the chart will automatically reconfigure to the new highest parameters.
  - DALY comparison. Similar to the previous bubble chart, x-axis is incidence per 100 000 and y-axis is DALYs per case.
  - Ranking results table. Final interactive ranking summary table: each row is a disease and columns include YLD, YLL, DALY, DALYs per case and DALYs per 100 000. When clicking on the heading of column, the ranking changes according to the ranking of the chosen column.
  - Ranking results bar chart. Expression of the previous table in a bar chart with percentile intervals.

Examples of results are reported in Appendix A.

3. Performance evaluation of the selected tools

The selected tools were evaluated using two case studies: a single pathogen–multiple food setting (*L. monocytogenes* in RTE foods) and multiple pathogens in a single food (STEC, *Salmonella* spp., *L. monocytogenes*, *Campylobacter* spp., Norovirus, *Cryptosporidium* spp. and *Giardia* spp.) in leafy greens. The P^3ARRT was not included in the evaluation since it was considered of the same structure with the EFoNAO-RRT.

As described in the previous section, the selected tools present significant differences in the risk metrics, the ranking approach, the model type, the model's variables and data integration. The objective of the evaluation of the tools in the two case studies performed in this section was to identify potential problems in using the tools and demonstrate the effect of the above differences in the risk ranking outputs. For this, the outputs of the different tools were also compared with a reference model developed by the BIOHAZ Panel. The reference model is a bottom-up, fully quantitative, stochastic, risk ranking model which follows the risk assessment paradigm and respects the laws of probability and calculus.

3.1. Development of a reference risk ranking model

The reference model is a retail-to-consumption model starting with the initial prevalence and concentration of the pathogens in the food products at retail. The growth of the pathogens during distribution and storage is calculated using the appropriate growth models based on the storage time and temperature. The concentration of the pathogen in a contaminated food at the time of consumption is calculated as the sum of the initial concentration and the growth during storage (in $\log_{10}$ scale). In order to take into account the maximum population density, an upper limit is set to the latter concentration. In the case of food products that are cooked before consumption, a decline of the pathogen during cooking is taken into account. The dose (cells per serving) is calculated as the product of the concentration at consumption time and the serving size using the Poisson distribution. The dose is then input to a dose–response model to calculate the probability of illness from the consumption of a contaminated serving. The probability of illness per serving is then estimated by multiplying the probability of illness per contaminated serving by the initial prevalence at retail. The total number of illnesses per annum is calculated as the product of the mean probability of illness per serving and the total number of servings per annum. Finally, the total number of servings per annum is multiplied by the DALYs per case to estimate the total DALYs.

The structure of the reference model is shown in detail in Figure 3 and Table 1.

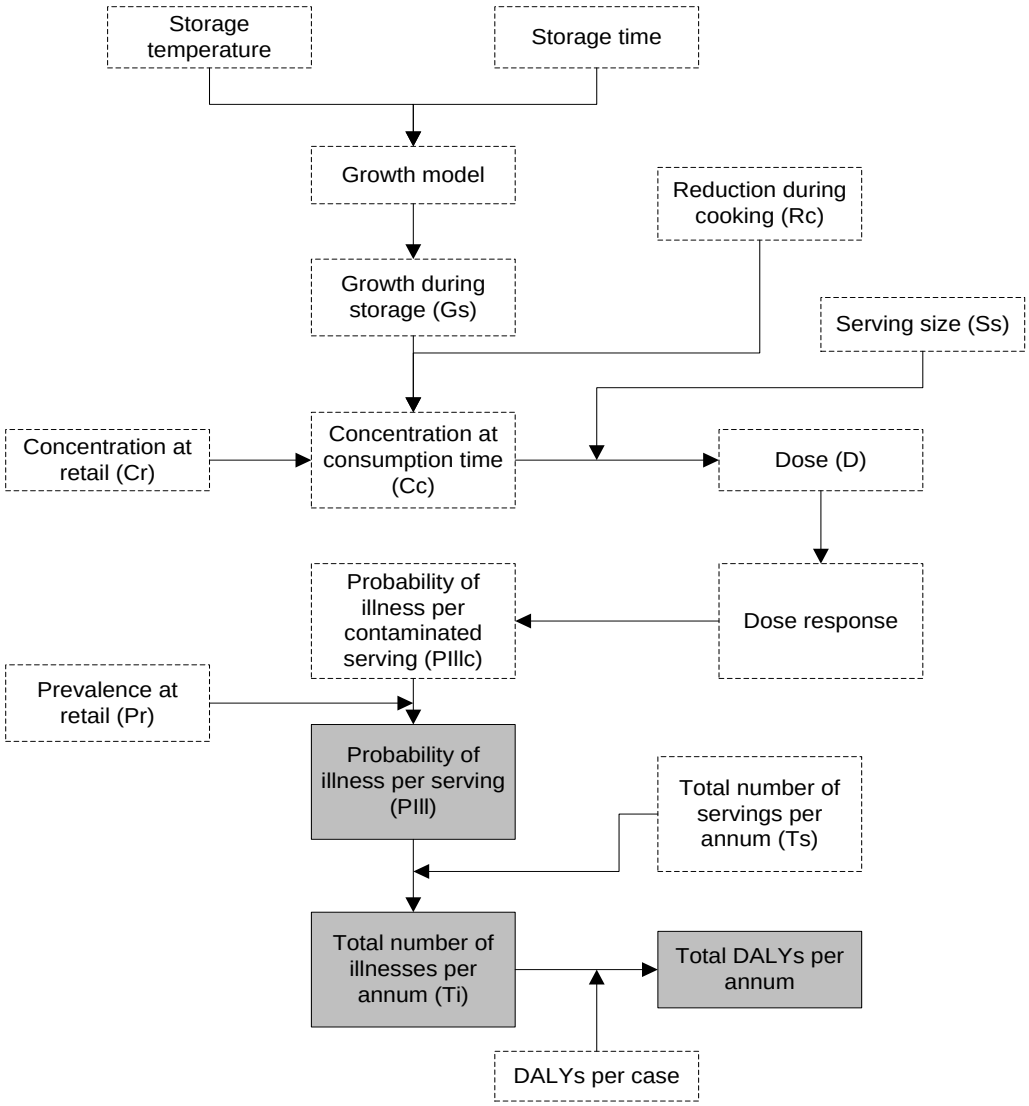


Figure 3: Structure of the reference model

Table 1: Variables and data integration used in the reference model

Input variable[(a)]	Description	Units	Data integration
Pr	Prevalence at retail	%	–
Cr	Concentration at retail	CFU/g	–
Gs	Growth during storage	CFU/g	Calculated from growth model based on the storage temperature and time
Rc	Reduction during cooking	CFU/g	–
Cc	Concentration at consumption time	Log_{10} CFU/g	log(Cr) + log(Gs) – log(Rc) with log(Cr) + log(Gs) ≤ log(Nmax)
Ss	Serving size	g	–
D	Dose (cells per serving)	CFU	Poisson (Cc × Ss)
PIllc	Probability of illness per contaminated serving	–	Calculated from dose–response model based on the dose
PIll	Probability of illness per serving	–	Pci × Pr
Ts	Total number of annual servings	–	–
Ti	Total number of illnesses per annum	–	Mean Pi × Ts
DALYs	DALYs per annum		Ti × DALYs per case

(a): Details of the variables are presented in the description of the databases of the two case studies.

In the reference model, only variability of the input variables was taken into account by introducing the input variables with probability distributions as presented in the description of the datasets of the two case studies. The model was run in Excel with @Risk using the Monte Carlo simulation with 30 000 iterations.

3.2. Application of the available tools to rank the risk of *L. monocytogenes* in selected RTE food categories

The available tools, decision trees, EFoNAO-RRT, Risk Ranger, microHibro, sQMRA, FDA-iRISK and BCoDE, were evaluated through an application exercise on risk ranking of *L. monocytogenes* in selected RTE food categories. The objective of the exercise was to apply the different tools using the same dataset, identify problems in using the tools and evaluate the performance of each tool based on specific criteria. The dataset used for the exercise was mainly based on the FDA/FSIS report on Quantitative assessment of the relative risk to public health from food-borne *L. monocytogenes* among selected categories of RTE foods (2003). The following five food categories were selected to be included in the exercise, representing different processing/storage conditions, consumer preparation, consumption patterns and risk at the time of consumption:

- smoked seafood
- soft ripened cheese
- pasteurised milk
- frankfurters (reheated)
- deli meats

Since the different tools require different input parameters, data extracted from FDA/FSIS report were modified in order to be applicable to all tested tools. The selection of data from the FDA/FSIS report does not aim to compare the outputs of the tested tools with that of the FDA/FSIS risk assessment model but to build a realistic database which would be applicable to all risk ranking tools.

The basic common dataset used for this case study is presented in detail in Table 2. The starting point of the exercise was the retail level. The database consists of 10 input parameters: (1) the prevalence of the pathogen at retail level; (2) the concentration of the pathogen at retail level; (3) the growth of the pathogen during domestic storage; (4) the reduction in the pathogen during consumer cooking in the case of frankfurters; (5) the serving size for each food category; (6) the total number of annual servings and (7) the population of interest (7) chosen for elderly population (more than 65 years of age); (8) the dose–response based on an exponential model; (9) the DALYs per case; (10) and the cost-of-illness per case As shown in Table 2, parameters 1 to 6 were different for each food category since they refer to the product, while parameters 7 to 10 were the same for all food categories since they refer to the consumer population or the pathogen.

Table 2: Common dataset used for the application exercise on risk ranking of *L. monocytogenes* in selected ready-to-eat food categories

Parameter	Smoked seafood	Soft ripened cheese	Pasteurised milk	Frankfurters (reheated)	Deli meats
1. Prevalence at retail (%)	6.4	1.5	0.3	5.5	7.5
2. Concentration at retail (arithmetic mean of contaminated product, CFU/g)	3 800	37	7	3 400 000	4 100
3. Growth at domestic storage (log_{10} CFU/g)	0.482	0.000	0.985	0.848	0.425
4. Reduction during cooking for frankfurters (log_{10} CFU/g)	–	–	–	3	–

Table 2: Common dataset used for the application exercise on risk ranking of *L. monocytogenes* in selected ready-to-eat food categories (continued)

Parameter	Smoked seafood	Soft ripened cheese	Pasteurised milk	Frankfurters (reheated)	Deli meats
5. Serving size (g)	61	35	228	76	60
6. Total number of annual servings for elderly	4.10E+07	1.80E+08	1.80E+10	5.80E+08	2.80E+09
7. Population of interest (elderly)	32 500 000	32 500 000	32 500 000	32 500 000	32 500 000
8. Dose–response (R of exponential model)	8.40E–12	8.40E–12	8.40E–12	8.40E–12	8.40E–12
9. DALYs/case	0.6	0.6	0.6	0.6	0.6
10. Cost-of-illness (€/case)	29 114	29 114	29 114	29 114	29 114

DALY: disability-adjusted life years.

The above parameters were used to test all tools to rank the risk of the chosen five food categories. Where necessary, parameters were translated according to the requirements of each tool. In addition, for the quantitative tools FDA-iRISK, sQMRA and microHibro, variability was taken into account for the input parameters described below.

Concentration of the pathogen at retail level

Data on the concentration of *L. monocytogenes* at retail from the FDA/FSIS report were fitted to the log-normal distribution. The parameters of the distribution and the mean concentration for the different food categories and the mean concentration are presented in Table 3.

Table 3: Parameter values of the log-normal distribution for the concentration of *L. monocytogenes* at retail

Food category	Log_{10} scale mean	Log_{10} scale SD	Arithmetic mean (CFU/g)
Smoked seafood	2.459	0.987	3 800
Soft-ripened cheese	1.152	0.601	37
Pasteurised milk	0.575	0.484	7
Frankfurters (reheated)	5.583	0.908	3 400 000
Deli meats	2.425	1.016	4 100

SD: standard deviation.

Growth of the pathogen during domestic storage

Growth rates during domestic storage were estimated using the cardinal model with inflection (CMI) originally developed by Rosso et al. (1993):

$$\mu_{max} = \frac{\mu_{max_{opt}}(T-T_{max})(T-T_{min})^2}{(T_{opt}-T_{min})[(T_{opt}-T_{min})(T-T_{opt})-(T_{opt}-T_{max})(T_{opt}+T_{min}-2T)]} \quad (1)$$

where T_{opt}, T_{min} and T_{max} (°C) are the theoretical optimum, minimum and maximum temperature for growth, respectively, and μ_{max} opt is the growth rate at optimum temperature. For *L. monocytogenes*, the following cardinal parameters were used as reported by Rosso et al. (1993): T_{opt} = 37 °C, T_{min} = 1.72 °C and T_{max} = 45.5 °C. Growth of the pathogen during domestic storage was estimated with an exponential primary model with no lag phase based on the parameters for storage temperature, storage time, optimum growth rate (expressed as minimum generation time, $GT_{min} = \log_2/\mu_{max}$) and maximum population density presented in Table 4 for the different food categories.

Table 4: Parameters used for the estimation of *L. monocytogenes* growth during domestic storage for the different food categories

Food category	Storage temperature (°C)	Storage time (hours) (exponential distribution)	Generation time at optimum temperature conditions	Maximum population density	Mean growth (log_{10} CFU/g)
Smoked seafood	6.5	Exp(96), max 720	2.69	6	0.482
Soft ripened cheese	6.5	0.39	Infinite	N/A	0.000
Pasteurised milk	6.5	Exp(96), max 288	1.11	9	0.985
Frankfurters (reheated)	6.5	Exp(120), max 359	1.62	8	0.848
Deli meats	6.5	Exp(120), max 360	3.23	6	0.425

N/A: not applicable.

The above approach was used for sQMRA and microHibro. In the case of FDA-iRISK, which does not provide the option of using a growth model, a custom probability for the total growth of the pathogens during domestic storage, estimated based on the approach below, using the Monte Carlo simulation (Figure 4) was used.

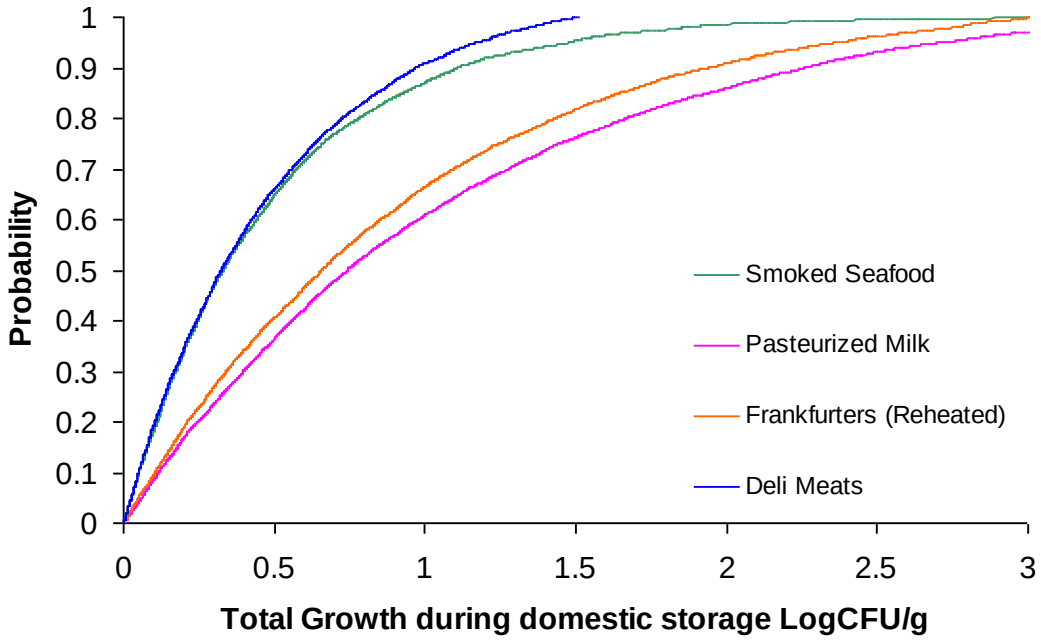


Figure 4: Custom probability for the total growth of *Listeria monocytogenes* during domestic storage

Serving size

Variability in serving size for FDA-iRISK, sQMRA and microHibro tools was described with a gamma distribution. The parameters of the distribution for the different food categories are presented in Table 5.

Table 5: Parameters of the gamma distribution used to describe the variability in the serving size for different food categories

Food category	Gamma distribution parameters		Mean serving size (g)
	a	b	
Smoked seafood	4.76	12.74	61
Soft ripened cheese	1.65	21.10	35
Pasteurised milk	2.99	76.20	228
Frankfurters (reheated)	1.37	55.6	76
Deli meats	4.83	12.40	60

3.2.1. Qualitative decision trees

Decision trees used in the opinion on public health risks represented by certain composite products (EFSA Panel on Biological Hazards (BIOHAZ), 2012a) were used for the examples considered in this Section. The decision trees were originally used to rank risks in certain composite products, based on food parameters impacting on growth/survival of the hazards involved, but were developed in order to be similarly applicable to all other foods. The one used for this specific example (i.e. *L. monocytogenes* in selected RTE food categories) is shown in Figure 2, Section 2.1.7, and relates to hazards that usually need to grow in food to cause illness.

3.2.1.1. Input parameters

The decision tree input parameters were selected based on the dataset presented in Table 3. The final input parameters for the five food categories are shown in Table 6.

Table 6: Input parameters of decision trees for the five food categories

1. Microbial treatment in package with no recontamination?	**Qualitative score**
Smoked seafood	No
Soft ripened cheese	No
Pasteurised milk	No
Frankfurters (reheated)	No
Deli meats	No
2. Supports growth?	
Smoked seafood	Yes
Soft ripened cheese	No
Pasteurised milk	Yes
Frankfurters (reheated)	Yes
Deli meats	Yes
3. Cooking before consumption?	
Smoked seafood	No
Soft-ripened cheese	N.A.
Pasteurised milk	No
Frankfurters (reheated)	Yes
Deli meats	No

N.A.: question not applicable for this food (negative answer to the previous question).

3.2.1.2. Risk ranking outputs

The ranking using the decision tree is shown in Table 7. The risk was qualified as low risk, moderate risk or QPR (Qualified Presumption of Risk) for *L. monocytogenes*. According to the decision tree, QPR means that the pathogen considered, if present, has the potential to cause disease via consumption of the food, and that the risk should be further qualified based on hygienic conditions in the preparation and/or on the possible growth of the pathogen before consumption, given the shelf life, storage temperature or conditions of use by the consumer (EFSA Panel on Biological Hazards

(BIOHAZ), 2012a). In this case, since it is considered that there is a possibility for growth of the pathogen before consumption, and that proper hygienic conditions may not be assumed, the risk for the three foods classified as QPR (smoked seafood, pasteurised milk and deli meats) has been further qualified as being high.

Table 7: Risk ranking outputs according to the decision tree for *Listeria monocytogenes* in selected ready-to-eat food categories

Product	Risk	Ranking
Smoked seafood	High	1
Soft-ripened cheese	Low	3
Pasteurised milk	High	1
Frankfurters (reheated)	Moderate	2
Deli meats	High	1

3.2.2. EFSA food of non-animal origin (EFoNAO)

3.2.2.1. Input parameters

Input data used for the evaluation are summarised in Table 8. In contrast to when EFoNAO-RRT was developed, this case study was not based on data from the EU. In the tool, category definitions were strictly based on EU data (e.g. for the epi-criteria). Here, US data were used instead. The CDC Foodborne outbreak online database was used to collect data on reported outbreaks and cases in the USA between 1998 and 2006 (CDC, online). Total cases and multipliers are based on Mead et al. (1999). However, the same category definitions as in the original tool were used. In the original tool, inactivation is not considered. This means that in the present evaluation of frankfurters, which commonly are reheated before consumption, only growth is considered and not inactivation, which may lead to an over-estimation of the risk. The scores for the criteria are presented in Table 9.

Table 8: Input data used for ranking of five ready-to-eat foods using the EFoNAO-RRT (EFSA BIOHAZ Panel, 2013)

Criterion	Input/data	Smoked seafood	Soft ripened cheese	Pasteurised milk	Frankfurters (reheated)	Deli meats	Comments
1. Epi-link	No of outbreaks	0	0	0	3[(a)]	7	CDC data[(b)]
	No of cases	0	0	0	109	142	
	Score	1	1	1	3	4	
2. Incidence	No of cases	1 259					
	Multiplier	2.0					
	Total cases[(c)]	2 493	2 493	2 493	2 493	2 493	
	Score	1	1	1	1	1	
3. Public health burden	DALYs per 1 000 cases	600	600	600	600	600	
	Score	3	3	3	3	3	
4. Dose–response	$IllD_{50}$ (log_{10} CFU) based on R=8.4E–12	10.92	10.92	10.92	10.92	10.92	
	Score	2	2	2	2	2	
5. Prevalence of contamination	Prevalence	6.4	1.5	0.3	5.5	7.5	
	Score	4	4	3	4	4	
6. Consumption	Percentage consuming	0.35	1.52	151	4.9	24	
	Score	1	2	4	3	3	

Table 8: Input data used for ranking of five ready-to-eat foods using the EFoNAO-RRT (EFSA BIOHAZ Panel, 2013) (continued)

Criterion	Input/data	Smoked seafood	Soft ripened cheese	Pasteurised milk	Frankfurters (reheated)	Deli meats	Comments
7. Growth potential	Growth during storage (log_{10}/g)	0.482	0	0.985	0.848	0.425	
	Growth score	4	1	4	4	4	
	Shelf life (days)	4 to 30	No growth	4 to 12	5 to 15	5 to 15	Estimated mean and maximum domestic storage times
	Shelf life Score	4	–	2	3	3	According to EFoNAO opinion scored according to longest shelf life

$IllD_{50}$: The dose needed to cause illness in 50 % of exposed humans.
(a): Used data for hot dogs. No Frankfurter outbreaks reported.
(b): The CDC Foodborne outbreak online database (http://wwwn.cdc.gov/foodborneoutbreaks/). Reported outbreaks in the USA between 1998 and 2006.
(c): Total number of cases as cited in Listeria FDA report (2003) was estimated by Mead et al. (1999), under the assumption of underreporting by a factor of 2 and that 99 % of cases are food borne. Estimated number of total cases using data in Scallan et al. (2011), under the assumption of under-reporting by a factor of 2.1, is 1 591 cases. This difference would not change the categorisation, i.e. the score.

3.2.2.2. Risk ranking outputs

Since the case study is of the type one pathogen and multiple foods, scores for the criteria linked only to the pathogen were the same for all foods and did not contribute to differences in total scores and thus, rank. Criteria related to the food and which contributed to differences in risk scores were epi-link, prevalence of contamination, consumption and growth potential. The epi-link explained the difference between deli meats and frankfurters, whereas consumption was a major contributor to the risk score of pasteurised milk and a third place ranking.

Table 9: Summary of criteria scores and total risk scores associated with *L. monocytogenes* in five ready-to-eat foods[(a)]

Criterion	Criterion number	Scores				
		Smoked seafood	Soft ripened cheese	Pasteurised milk	Frankfurters (reheated)	Deli meats
Epi-link	1	1	1	1	3	4
Incidence	2	1	1	1	1	1
Public health burden	3	3	3	3	3	3
Dose–response	4	2	2	2	2	2
Prevalence of contamination	5	4	4	3	4	4
Consumption	6	1	2	4	3	3
Growth potential	7	4	1	3	4	4

(a): Scores were estimated by the EFoNAO risk ranking tool using input data in Table 2 and equal weights for all criteria.

Table 10: Risk ranking outputs of EFSA food of non-animal origin for *L. monocytogenes* in selected ready-to-eat food categories

Product	Total score	Ranking
Smoked seafood	16	4
Soft ripened cheese	14	5
Pasteurised milk	17	3
Frankfurters (reheated)	20	2
Deli meats	21	1

3.2.3. Risk Ranger

3.2.3.1. Input parameters

The input parameters of Risk Ranger tool were selected based on the dataset presented in Table 2. For some parameters, the options provided by Risk Ranger for the values did not match with the dataset. In this case, the option with the closest value to the dataset was selected. The final input parameters for the five food categories are shown in Table 11.

Table 11: Input parameters of Risk Ranger for the five food categories

1. Hazard severity	Score	Numerical[(a)]
Smoked seafood	MODERATE hazard	0.01
Soft ripened cheese	MODERATE hazard	0.01
Pasteurised milk	MODERATE hazard	0.01
Frankfurters (reheated)	MODERATE hazard	0.01
Deli meats	MODERATE hazard	0.01
2. How susceptible is the consumer?		
Smoked seafood	GENERAL	1
Soft ripened cheese	GENERAL	1
Pasteurised milk	GENERAL	1
Frankfurters (reheated)	GENERAL	1
Deli meats	GENERAL	1
3. Frequency of contamination		
Smoked seafood	Common (50 %)	0.5
Soft ripened cheese	Common (50 %)	0.5
Pasteurised milk	Common (50 %)	0.5
Frankfurters (reheated)	Common (50 %)	0.5
Deli meats	Common (50 %)	0.5
4a. Effect of process		
Smoked seafood	The process RELIABLY ELIMINATES hazards	0
Soft ripened cheese	The process RELIABLY ELIMINATES hazards	0
Pasteurised milk	The process RELIABLY ELIMINATES hazards	0
Frankfurters (reheated)	The process RELIABLY ELIMINATES hazards	0
Deli meats	The process RELIABLY ELIMINATES hazards	0
4b. Effect of preparation for meals		
Smoked seafood	Meal Preparation has NO EFFECT on the hazards	1
Soft ripened cheese	Meal Preparation has NO EFFECT on the hazards	1
Pasteurised milk	Meal Preparation has NO EFFECT on the hazards	1
Frankfurters (reheated)	OTHER	1.00E–03
Deli meats	Meal preparation has NO EFFECT on the hazards	1

Table 11: Input parameters of Risk Ranger for the five food categories (continued)

5. Is there potential for recontamination?		
Smoked seafood	OTHER	0.064
Soft ripened cheese	OTHER	0.015
Pasteurised milk	OTHER	0.003
Frankfurters (reheated)	OTHER	0.055
Deli meats	OTHER	0.075
6. How effective is the post-processing control system?		
Smoked seafood	NOT CONTROLLED	3.00
Soft ripened cheese	WELL CONTROLLED	0.00
Pasteurised milk	NOT CONTROLLED	10.00
Frankfurters (reheated)	NOT CONTROLLED	10.00
Deli meats	NOT CONTROLLED	3.00
7. How much increase is required to reach an infectious or toxic dose?		
Smoked seafood	OTHER	2.83E+06
Soft ripened cheese	OTHER	5.05E+08
Pasteurised milk	OTHER	2.00E+06
Frankfurters (reheated)	OTHER	3.10E+08
Deli meats	OTHER	3.75E+05
8. Frequency of consumption		
Smoked seafood	A few times per year	3
Soft ripened cheese	A few times per year	3
Pasteurised milk	Daily	365
Frankfurters (reheated)	Monthly	12
Deli meats	Weekly	52
9. Proportion of consuming population		
Smoked seafood	All (100 %)	1
Soft ripened cheese	All (100 %)	1
Pasteurised milk	All (100 %)	1
Frankfurters (reheated)	All (100 %)	1
Deli meats	All (100 %)	1
10. Size of consuming population		
Smoked seafood	OTHER	32 500 000
Soft ripened cheese	OTHER	32 500 000
Pasteurised milk	OTHER	32 500 000
Frankfurters (reheated)	OTHER	32 500 000
Deli meats	OTHER	32 500 000

(a): See Section 2.4.2 for description of risk metrics of the tool.

3.2.3.2. Risk ranking output

The ranking of the three risk metrics provided by Risk Ranger is shown in Table 12. The ranking output was the same for the probability of illness per day per consumer of interest, the total predicted illnesses/annum in population of interest and the risk ranking metrics. The tool ranked the foods in the following order of decreasing risk: deli meats, pasteurised milk, smoked seafood, soft ripened cheese and frankfurters.

Table 12: Risk ranking outputs of Risk Ranger for *L. monocytogenes* in selected ready-to-eat food categories

Product	Probability of illness per day per consumer of interest	Ranking
Smoked seafood	1.86E–09	3
Soft ripened cheese	2.44E–13	4
Pasteurised milk	1.50E–08	2
Frankfurters (reheated)	5.84E–14	5
Deli meats	3.18E–07	1
Product	**Total predicted illnesses/annum in population of interest**	**Ranking**
Smoked seafood	22.04	3
Soft ripened cheese	0.003	4
Pasteurised milk	177.7	2
Frankfurters (reheated)	0.001	5
Deli meats	3 767	1
Product	**Risk ranking**	**Ranking**
Smoked seafood	39	3
Soft ripened cheese	17	4
Pasteurised milk	44	2
Frankfurters (reheated)	13	5
Deli meats	52	1

3.2.4. microHibro

3.2.4.1. Input parameters

The input parameters of microHibro were selected based on the dataset presented in Table 2, modified as described below. Parameters can be entered as fixed or variable. For variable inputs, microHibro proposes a limited choice of distributions (normal, gamma, beta, exponential, uniform, triangular), as it also does for discrete ones (binomial and Poisson), although the tool allows the inclusion of new ones.

For prevalence and concentration at retail microHibro can use the distributions mentioned above, which can be truncated to a maximum and minimum values. The food portion size can be introduced and taken into account in the calculations. For growth data, microHibro presents a selection of published growth models for different purposes that can be selected. In addition, the user can also introduce new growth models in the application and use them for the calculations. For inactivation, it can work with direct input of log-reduction fixed values or distributions. Consumption can be described by the distributions indicated; in this case study the gamma distribution was used. Eating occasions can be implemented as a fixed value.

For the probability of illness, microHibro uses a log scale in the calculation so it actually provides the mean of the log probability of illness. Dose–response is calculated as r-value, but alternative models can be included by the user and selected for the calculations.

3.2.4.2. Risk ranking outputs

The outputs from microHibro are presented in Table 13.

Table 13: Risk ranking stochastic outputs of microHibro for *L. monocytogenes* in selected RTE food categories

Product	Mean probability of illness per day per consumer of interest	Ranking
Smoked seafood	2.89E–09	2
Soft ripened cheese	3.77E–10	4
Pasteurised milk	2.69E–09	3
Frankfurters (reheated)	6.16E–13	5
Deli meats	8.38E–09	1

3.2.5. Swift quantitative microbiological risk assessment (sQMRA)

3.2.5.1. Input parameters

Not all the inputs specified in Table 2 can be entered directly into the sQMRA model; some pre-processing of inputs is necessary as the tool only accepts inputs in one format (see Table 14). In some cases, pre-processing is straightforward, e.g. calculating the consumption in portions per month from the total population and the annual consumption, or specifying portions sizes by the mean and standard deviation and not by the parameters of the underlying gamma distribution. In other cases, more effort is needed and may require considerable skills in quantitative microbiology, e.g. knowledge of growth models. sQMRA cannot accept direct input of log-growth but calculates growth according to an exponential growth model with a gamma model for the impact of temperature on the growth rate. It was assumed that all food is stored in the fridge. Cardinal growth parameters for *L. monocytogenes* were taken from Augustin et al. (2005). Storage temperature was then empirically adjusted to achieve average log-growth in the deterministic model as specified for the different products in Table 8.

For soft ripened cheese, storage temperature was set to –2 °C to force the model into die-off mode. The maximum population density was set at 10^6 CFU/g for all food products except for frankfurters, for which it was set at 10^8.

Mean storage time was used as the average of the most likely range in the original report, while maximum storage time was used as the maximum of the maximum range. For frankfurters, the average storage time was calculated as the sum product of the full distribution; the maximum was set at 21 days to prevent extremely skewed distributions.

Table 14: Input parameters specific to the sQMRA models

Input parameter	Comments
Prevalence at retail	sQMRA models only variability
Concentration at retail	sQMRA cannot work with percentiles, only with log-normal distributions. Output from the FDA model is highly skewed, and a log-normal distribution does not adequately fit these data. Only average was used.
Growth during domestic storage	sQMRA cannot accept direct input of log-growth but calculates growth according to an exponential growth model with a gamma model for the impact of temperature on the growth rate. It was assumed that all food is stored in the fridge. Mean storage time was used as the average of the most likely range, while maximum storage time was used as the maximum of the maximum range. For frankfurters, average storage time was calculated as the sum product of the full distribution; the maximum was set at 21 days to prevent extremely skewed distributions. Cardinal growth parameters were taken from Augustin et al. (2005). T_{min} = –1.72 °C; T_{opt} = 37 °C. μ_{opt} (h^{-1} at 37 ºC) in smoked seafood: 0.549; soft ripened cheese: 0.000; pasteurised milk: 0.941; frankfurters (reheated): 0.480; deli meats: 1.033. Storage temperature was empirically adjusted to achieve average log-growth as specified for the different products.

Table 14: Input parameters specific to the sQMRA models (continued)

Input parameter	Comments
Growth during domestic storage (continued)	For soft ripened cheese, storage temperature was set to –2 °C to force the model into die-off mode. Maximum population density set at 1E6; 1E8 for frankfurters.
Reduction during cooking	Probability of survival calculated as 10^(–log-reduction). Only average value used as using also 5th and 95th percentile has a major impact on the average reduction.
Serving size	Only mean was used
Consumption	Population size is 3.24E7; 13 % of total population specified in spreadsheet. sQMRA requires number of servings per person-months; this was adjusted to achieve the specified number of serving annually.
DALYs/cost-of-illness	Direct inputs as specified
Dose–response	r-value for dose–illness model as specified.

sQMRA: swift quantitative microbiological risk assessment.

3.2.5.2. Risk ranking outputs

The outputs obtained after running the sQMRA tool in both the deterministic and stochastic approaches are presented in Tables 15 and 16.

Table 15: Risk ranking outputs of sQMRA for *L. monocytogenes* in selected ready-to-eat food categories, using the deterministic approach

Product	Mean probability of illness per day per consumer of interest	Ranking
Smoked seafood	4.76E–08	2
Soft ripened cheese	2.06E–11	4
Pasteurised milk	1.00E–08	3
Frankfurters (reheated)	8.68E–13	5
Deli meats	3.69E–07	1
Product	**Total predicted illnesses/annum in population of interest**	**Ranking**
Smoked seafood	1.95	3
Soft ripened cheese	0.004	4
Pasteurised milk	180	2
Frankfurters (reheated)	0.001	5
Deli meats	1 033	1
Product	**DALYs**	**Ranking**
Smoked seafood	1.2	3
Soft ripened cheese	0.002	4
Pasteurised milk	108	2
Frankfurters (reheated)	0.001	5
Deli meats	620	1

sQMRA: swift quantitative microbiological risk assessment.

Table 16: Risk ranking outputs of sQMRA for *L. monocytogenes* in selected ready-to-eat food categories, using the stochastic approach

Product	**Mean probability of illness per day per consumer of interest**	**Ranking**
Smoked seafood	9.6E–08	2
Soft ripened cheese	2.1E–11	4
Pasteurised milk	2.7E–08	3
Frankfurters (reheated)	8.8E–12	5
Deli meats	1.1E–07	1
Product	**Total predicted illnesses/annum in population of interest**	**Ranking**
Smoked seafood	4	3
Soft ripened cheese	0.004	5
Pasteurised milk	500	1
Frankfurters (reheated)	0.005	4
Deli meats	307	2
Product	**DALYs**	**Ranking**
Smoked seafood	2.4	3
Soft ripened cheese	0.002	5
Pasteurised milk	300	1
Frankfurters (reheated)	0.003	4
Deli meats	184	2

sQMRA: swift quantitative microbiological risk assessment.

The output results of sQMRA are provided at a very detailed level, for different steps in the food chain. They are therefore very useful to evaluate the impact of using different risk metrics for ranking purposes. For a single pathogen in multiple food products, including DALYs and cost-of-illness as risk metrics does not affect the ranking; nevertheless, comparing the ranking results for different metrics provides important insights.

3.2.6. FDA-iRISK

3.2.6.1. Input parameter

The input parameters of FDA-iRISK were selected based on the dataset presented in Table 17. The input parameters in FDA-iRISK can be entered as fixed or variable. When the input has to be considered as variable, FDA-iRISK proposes a limited choice of distribution: beta-PERT, empirical, normal, triangular and uniform. For some parameters, the options provided by FDA-iRISK for the distribution did not match with the dataset. In this case empirical distribution was selected.

Table 17: Input parameters specific to the FDA-iRISK model

Input parameter	Comments
Prevalence at retail	FDA-iRISK use only fixed values of prevalence
Concentration at retail	FDA-iRISK can work with parametric distributions (beta-PERT, normal, triangular and uniform) and non-parametric distribution: cumulative empirical distribution For the initial concentration, it was assumed that the concentrations of log_{10} CFU/g follow normal distribution with the parameters:

Foods	Mean	SD
Smoked seafood	2.46	0.987
Soft ripened cheese	1.15	0.601
Pasteurised milk	0.57	0.48
Frankfurters (reheated)	5.58	0.91
Deli meats	2.42	1.02

Growth during domestic storage	FDA-iRISK accepts direct input of log-growth fixed values or distributions The growth was first calculated according to an exponential growth model with a gamma model for the impact of temperature on the growth rate (as described in Table 14 in sQMRA inputs parameters). As the temperature and duration of storage vary between consumers, a Monte Carlo simulation model was run to obtain a cumulative empirical distributions of log_{10} growth (Figure 4)
Reduction during cooking	FDA-iRISK accepts direct input of log-reduction fixed values or distributions We used a cumulative empirical distribution to describe the variability of log-reduction during cooking of frankfurters
Consumption	***Portion size***: As gamma distribution is not implemented in FDA-iRISK, the cumulative empirical distributions of the gamma distributions with parameter a and b were first calculated (see Table 5) ***Eating occasions*** per year are fixed values in FDA-iRISK
DALYs/cost-of-illness	Direct inputs as specified
Dose–response	r-value for dose–illness model as specified

DALY: disability-adjusted life years; SD: standard deviation; sQMRA: swift quantitative microbiological risk assessment.

3.2.6.2. Risk ranking outputs

Table 18 presents the FDA-iRISK output results using the deterministic approach. The ranking order obtained with FDA-iRISK using the DALY metrics was: deli meats, pasteurised milk, smoked seafood, soft ripened cheese and frankfurters.

Table 18: Risk ranking outputs of FDA-iRISK for *L. monocytogenes* in selected ready-to-eat food categories using the deterministic approach

Product	Mean probability of illness per day per consumer of interest	Ranking
Smoked seafood	4.76E–08	2
Soft ripened cheese	2.06E–11	4
Pasteurised milk	1.00E–08	3
Frankfurters (reheated)	8.68E–13	5
Deli meats	3.69E–07	1

Table 18: Risk ranking outputs of FDA-iRISK for *L. monocytogenes* in selected ready-to-eat food categories using the deterministic approach (continued)

Product	Total predicted illnesses/annum in population of interest	Ranking
Smoked seafood	1.95	3
Soft ripened cheese	0.004	4
Pasteurised milk	180	2
Frankfurters (reheated)	0.001	5
Deli meats	1 033	1
Product	**DALYs**	**Ranking**
Smoked seafood	1.2	3
Soft ripened cheese	0.002	4
Pasteurised milk	108	2
Frankfurters (reheated)	0.001	5
Deli meats	620	1

DALY: disability-adjusted life years.

Table 19 presents the FDA-iRISK output results using the stochastic approach. The order obtained with FDA-iRISK using the DALY metrics was pasteurised milk, deli meats, smoked seafood, frankfurters and soft ripened cheese.

Table 19: Risk ranking outputs of FDA-iRISK for *L. monocytogenes* in selected ready-to-eat food categories using the stochastic approach

Product	Mean probability of illness per day per consumer of interest	Ranking
Smoked seafood	6.25E–06	1
Soft ripened cheese	2.06E–11	5
Pasteurised milk	5.10E–07	2
Frankfurters (reheated)	1.40E–10	4
Deli meats	3.15E–07	3
Product	**Total predicted illnesses/annum in population of interest**	**Ranking**
Smoked seafood	256	3
Soft ripened cheese	0.004	5
Pasteurised milk	9 180	1
Frankfurters (reheated)	0.081	4
Deli meats	882	2
Product	**DALYs**	**Ranking**
Smoked seafood	154	3
Soft ripened cheese	0.002	5
Pasteurised milk	5 508	1
Frankfurters (reheated)	0.049	4
Deli meats	529	2

DALY: disability-adjusted life years.

3.2.7. Burden of Communicable Diseases in Europe (BCoDE)

3.2.7.1. Input parameters

There are currently no attribution data at the EU level for the proportion of listeriosis that is food borne, or the foods associated with food-borne listeriosis. Therefore, using BCoDE in a top-down approach is not currently possible. Therefore, the BCoDE toolkit was used in combination with a bottom-up tool, i.e. the number of cases as predicted by the sQMRA model, currently considered as the tool that most precisely reflects the outputs of an unconstrained model.

The following adaptations from the sQMRA output were made: study population was considered to be men and women ≥ 65 years of age (13 % of overall population, 32 400 000), which we redistributed to the five age categories according to the EU population, for both men and women; the numbers of cases resulting from the sQMRA output (total predicted illnesses/annum) were distributed to the TESSy-notified cases of listeriosis in men and women ≥ 65 year of age (five age groups for each sex). The other main denominator, life expectancy, remained the same.

We input in the BCoDE toolkit the resulting incidence tables and set for 1 000 iterations; hence, the Monte Carlo simulation was run 1 000 times.

3.2.7.2. Risk ranking output

The ranking of the BCoDE toolkit is based on the number of cases of listeriosis as a result of the different foods. Absolute amount of DALYs, DALYs per 100 000, YLD and YLL per 100 000 can be expressed as a mean and median, as well as uncertainty intervals (2.5^{th} and 97.5^{th} percentiles). The following outputs are based on the ≥ 65 years of age population and are the median results only (Table 20).

Table 20: Risk ranking outputs according to the BCoDE for *L. monocytogenes* in selected ready-to-eat food categories

Product	**DALYs per 100 000**	**Ranking**
Smoked seafood	0.04	3
Soft ripened cheese	3.45289E–05	5
Pasteurised milk	4.58	1
Frankfurters (reheated)	4.67141E–05	4
Deli meats	2.81	2
Product	**DALYs**	**Ranking**
Smoked seafood	11.9	3
Soft ripened cheese	0.011	5
Pasteurised milk	1 483	1
Frankfurters (reheated)	0.015	4
Deli meats	910	2
Product	**YLD per 100 000**	**Ranking**
Smoked seafood	0.002	3
Soft ripened cheese	2.03E–06	5
Pasteurised milk	0.27	1
Frankfurters (reheated)	2.73E–06	4
Deli meats	0.16	2
Product	**YLL per 100 000**	**Ranking**
Smoked seafood	0.03	3
Soft ripened cheese	3.25E–05	5
Pasteurised milk	4.31	1
Frankfurters (reheated)	4.40E–05	4
Deli meats	2.64	2

DALY: disability-adjusted life years; YLD: years lived with disability; YLL: years of life lost as a result of premature mortality.

3.2.8. Comparison of the different tools

3.2.8.1. Comparison of risk metrics

The overall results of the reference risk assessment model and the different risk ranking tools for the *L. monocytogenes* case study in RTE foods are presented in Table 22. As shown in Table 22, there are significant deviations between the outputs of the different tools and the baseline models as well as among the tools. This can be attributed to the differences among the tools described in Section 2 related to the risk metrics, the ranking approach, the model type, the model variables and data

integration method. In the following paragraphs, a comparative analysis is performed in order to identify the sources of the differences between the outputs of the different tools and the baseline models.

A comparison in the probability of illness per serving estimated from the reference model and the bottom-up tools FDA-iRISK, sQMRA, microHibro and Risk Ranger is presented in Figure 5.

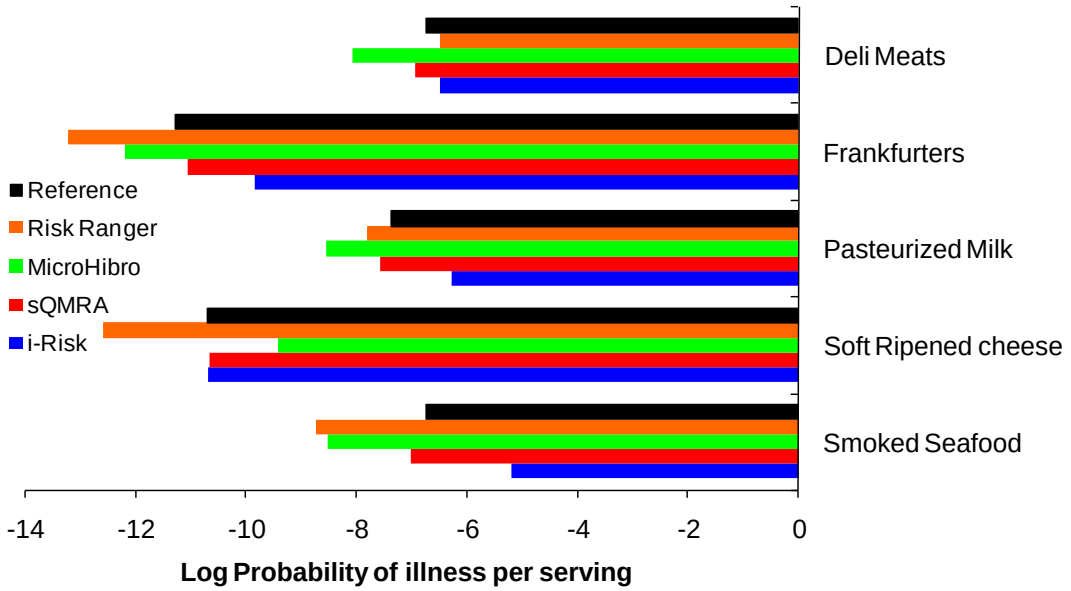


Figure 5: Comparison between the mean probabilities of illness per serving estimated from the reference model and the bottom-up tools FDA-iRISK, sQMRA, microHibro and Risk Ranger

The **FDA-iRISK** provided, in general, higher probabilities of illness per serving than the baseline model for all tested products that, according to the dataset used, support growth during storage (i.e. smoked seafood, pasteurised milk, frankfurters and deli meats). The higher predicted probabilities from FDA-iRISK can be attributed to the fact that the current version of the tool does not take into account a maximum population density of the pathogen. As a result, the summation of the initial concentration and the growth during storage may result in unrealistically high concentrations of the pathogen at the time of consumption and thus to higher probability of illness per serving. This can be seen in the case of soft ripened cheese in which *Listeria* cannot grow and the output of FDA-iRISK is identical to that of the baseline model since the maximum population density is less important.

The outputs of **sQMRA** were the closest to the reference model. This shows that the tool includes all the main factors affecting risk and follows the risk assessment paradigm respecting the laws of probability and calculus.

Significant deviations were observed between the outputs of **microHibro** and the reference model. These differences can be mainly attributed to the calculations of the mean probabilities of illness per serving. microHibro uses a log scale in the calculation so it actually provides the mean of the log probability of illness, which can be significantly different from the arithmetic mean of the probability of illness. Another source of these deviations is the fact that, owing to the Monte Carlo process that is run within the tool, in microHibro a small number of iterations was used.

Risk Ranger is also based on the bottom-up approach but the model type is deterministic. Risk Ranger provided, in general, lower values for the probability of illness per serving than did the reference model. This is mainly because this tool uses single values of the input parameters and does not take into account their variability. Another reason is that Risk Ranger does not use a full dose–

response relationship. Instead, a threshold value is assumed for the contamination level that would cause infection or intoxication to the average consumer, without taking into account variability in the dose–response. In addition, for some input parameters, the options for their value provided in the risk spreadsheet are limited. In this case, the offered option with the closest value to input data was selected.

Figure 6 presents a comparison between the deterministic outputs of the reference model and FDA-iRISK, sQMRA and Risk Ranger. microHibro is not included in the comparison because the tool cannot take into account initial prevalence in deterministic mode. The deterministic outputs of FDA-iRISK and sQMRA were identical, with the baseline models indicating that the differences are associated with incorporation of the variability of input parameters. The outputs of Risk Ranger were still different from the baseline model for the reasons explained above but the deviations were smaller than in the stochastic baseline model.

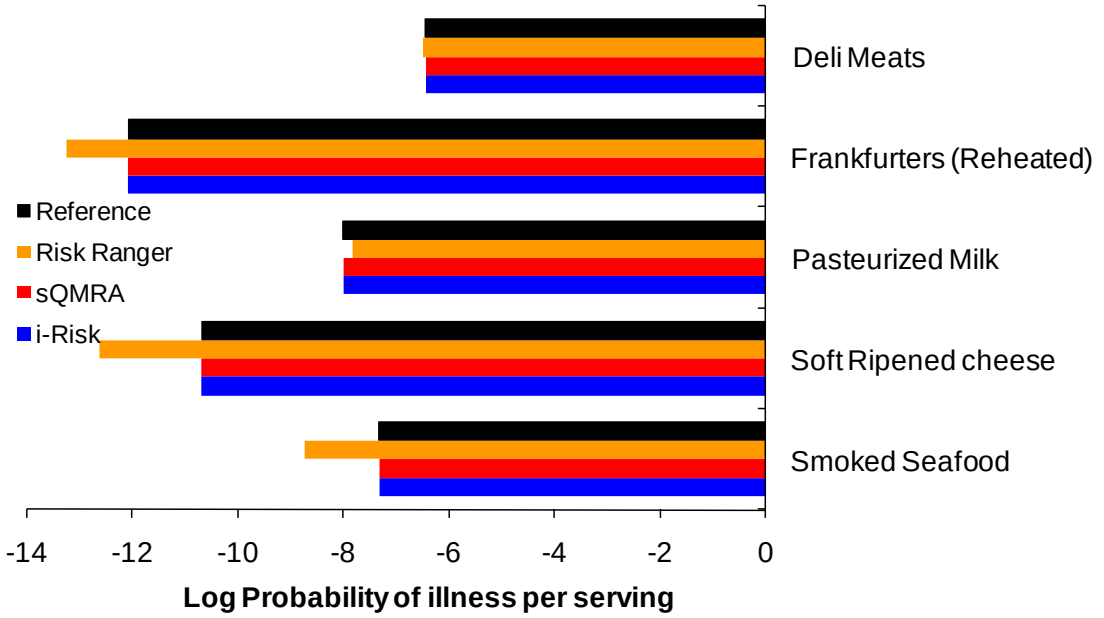


Figure 6: Comparison of the deterministic estimation of the probability of illness per serving between the baseline model the Risk Ranger, FDA-iRISK and sQMRA

EFoNAO and decision tree tools provide ordinal or categorical risk metrics and thus cannot be compared with the other tools.

In this case study, BCoDE was applied as a DALY calculator using the probability of illness estimated by the sQMRA as input parameter. As shown in Table 22, the DALY outputs of BCoDE were similar but not identical to those estimated by sQMRA, reflecting the different approach in DALY estimation.

3.2.8.2. Comparison of risk rankings

The output of the tested tools can be used to rank the risk and compare the ranking for the different tools. However, risk ranking requires a selection of a risk metric for each tool. The results of the present case study showed that even for the same tool the risk ranking can differ significantly for different risk metrics. An example is shown in Table 21, where the risk ranking from FDA-iRISK and sQMRA based on the mean probability of illness and total predicted illnesses/annum in population of interest are presented. For both tools the ranking changed for the different risk metrics.

Table 21: Risk ranking from FDA-iRISK and sQMRA based on the mean probability of illness and total predicted illnesses/annum in population of interest

	FDA-iRISK		sQMRA	
	Probability of illness/serving	**Total illnesses**	**Probability of illness/serving**	**Total illnesses**
Smoked seafood	2	3	2	3
Soft ripened cheese	4	4	4	5
Pasteurised milk	3	2	3	1
Frankfurters	5	5	5	4
Deli meats	1	1	1	2

An overall comparison of the rankings provided by the different tools is presented in Figure 7. The rankings are based on DALYs for FDA-iRISK, sQMRA, microHibro, Risk Ranger and BCoDE, on the sum of scores for EFoNAO and on the categorisation of risk for the decision tree. The figure shows clearly that the ranking is significantly affected by the ranking approach, the model type, the model variables and data integration method.

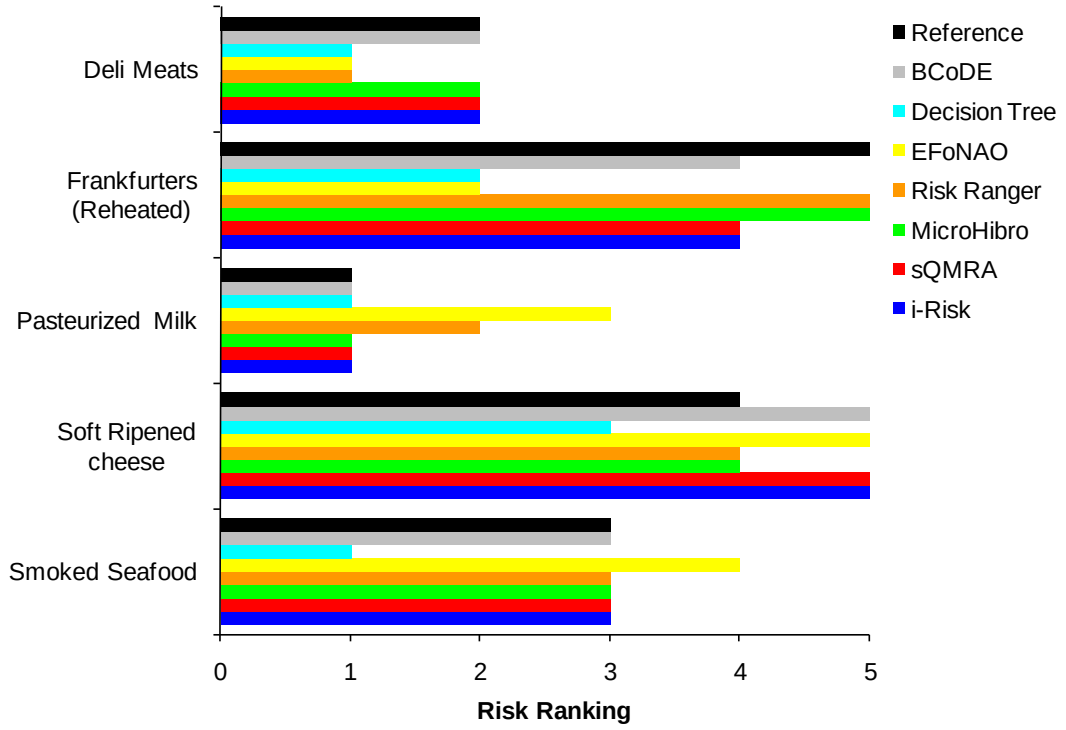


Figure 7: Overall comparison of the rankings provided by the different tools

Table 22: Output overview of the tested tools for the risk ranking of *L. monocytogenes* in selected ready-to-eat food categories

Tool/risk metric	Product				
	Smoked seafood	**Soft-ripened cheese**	**Pasteurised milk**	**Frankfurters (reheated)**	**Deli meats**
Reference model					
Probability of illness/serving	1.83–07	2.05E–11	4.15E–08	5.15E–12	1.88E–07
Total illnesses	7.51	0.004	747	0.003	526
DALYs	4.51	0.002	448	0.002	316
Risk Ranger					
Probability of illness/serving	1.86E–09	2.44E–13	1.50E–08	5.84E–14	3.18E–07
Total illnesses	22	0.003	178	0.001	3 767
DALYs[(a)]	13.2	0.002	106.8	0.0006	2.26
FDA-iRISK					
Probability of illness/serving	6.25E–06	2.06E–11	5.10E–07	1.40E–10	3.15E–07
Total illnesses	256	0.004	9 180	0.08	882
DALYs	154	0.002	5 508	0.05	529
sQMRA					
Probability of illness/serving	9.60E–08	2.1E–11	2.70E–08	8.80E–12	1.10E–07
Total illnesses	4	0.004	500	0.005	307
DALYs	2.4	0.002	300	0.003	184
EFoNAO					
Total Score	16	14	17	20	21
microHibro					
Probability of illness/serving	2.89E–09	3.77E–10	2.69E–09	6.16E–13	8.38E–09
Total illnesses	0.12	0.07	48	0.0004	23.5
DALYs	0.07	0.04	29	0.0002	14.1
Decision tree					
Risk evaluation	High	Low	High	Moderate	High
BCoDE					
DALYs	11.87	0.01	1 483	0.015	910

DALY: disability-adjusted life years.
(a): DALYs were estimated manually by multiplying the total number of illnesses per annum by the DALYs per illness case.

3.3. Application of the available tools to rank the risk of multiple pathogens in leafy greens

3.3.1. Input parameters used

The available tools FDA-iRISK, Risk Ranger, sQMRA, EFoNAO-RRT, decision trees and BCoDE were also evaluated through a second case study on risk ranking of multiple pathogens in leafy greens. The objective of the exercise was again to evaluate the different tools using the same dataset, identify problems in using and evaluate the performance of each tool using specific criteria. The following seven pathogens were considered relevant for leafy greens and included in the case study:

- STEC
- *Salmonella*
- *L. monocytogenes*
- *Campylobacter*
- Norovirus
- *Cryptosporidium*
- *Giardia*

Available data from the literature were collected to generate a common dataset for all tools. As in the case of the *Listeria* case study, the objective of this exercise was not to assess the risk but to compare the different tools using a common realistic dataset. These data are presented in detail in Tables 23 to 27.

Table 23: Initial prevalence and concentration of pathogens in leafy greens (derived from Robertson and Gjerde, 2001; Baert et al., 2011; Wijnands et al., 2014)

Pathogen	Prevalence (%)	Initial concentration (CFU/g)
STEC	0.54	0.052
Salmonella	0.17	0.024
L. monocytogenes	1.77	250
Campylobacter	0.083	0.024
Norovirus	0.165	100
Cryptosporidium	4	0.03
Giardia	2	0.025

The above parameters were used as input in the tools evaluated to rank the risk of the seven pathogens. Where necessary, parameters were translated according to the requirements of each tool. In addition, for the quantitative tools FDA-iRISK, sQMRA and microHibro, variability was taken into account for the following input parameters (Tables 24 to 27).

Table 24: Cardinal model parameters for the growth of STEC, *Salmonella* and *L. monocytogenes* during storage (derived from Rosso et al., 1993; Koseki and Isobe, 2005a, b). For the rest of pathogens no change in the concentration during storage was assumed.

Cardinal parameters	STEC	*Salmonella*	*L. monocytogenes*
T_{min}	4.9	5.7	1.72
T_{max}	41.3	40	37
T_{opt}	47.5	49.3	45.5
m_{opt}	2.5	1.96	0.76

Table 25: Cumulative probability of the storage time for leafy greens (Marklinder et al., 2004)

Storage time (days)	Probability
0	0
1	0.61
2	0.88
3	0.93
7	1

The storage temperature was defined with a gamma distribution with the following parameters: a = 7.15, b = 1.03, min = 1.8, max = 18.2.

Table 26: Cumulative probability of serving size for leafy greens (derived from Carrasco et al., 2010)

Serving size (g)	Probability
25	0
28	0.5
55	0.75
123	0.95
200	1

Table 27: Dose–response parameters used in the leafy greens case study

Hazards	Type	p1	p2	Dose response type	Probability (illness| infection)	References
Campylobacter	Beta Poisson	1.45E–01	7.59E+00	Infection	33 %	(FAO/WHO, 2009)
Cryptosporidium	Exponential	5.73E–02		Infection	10 %	(Teunis et al., 2002)
Giardia	Exponential	1.99E–02		Infection	10 %	(Teunis et al., 1996)
L. monocytogenes	Exponential	8.40E–12		Illness		(U.S. FDA, 2003)
Norovirus	Exponential	5.00E–01		Illness	10 %	(Teunis et al., 2008)
Salmonella	Beta Poisson	1.32E–01	5.15E+01	Illness		(FAO/WHO, 2002)
STEC	Exponential	1.13E–03		Illness		(Strachan et al., 2005)

p1 = alpha; p2 = beta

3.3.2. Qualitative decision trees

Two decision trees from the opinion on public health risks represented by certain composite products (EFSA Panel on Biological Hazards (BIOHAZ), 2012a) were used for this case study: (1) the decision tree related to hazards which usually need to grow in food to cause illness (used for the first case study, see Figure 2 and Section 3.2.1) and (2) the decision tree related to hazards which may not need to grow in food to cause illness. In this case study, the former was used for *L. monocytogenes*, while the latter was used for the other pathogens. Compared with the decision tree showed in Figure 2, the second decision tree does not include a question related to the ability of the food to support the growth of the pathogen, as this is not an important parameter to consider for these pathogens.

3.3.2.1. Input parameters

The decision tree input parameters were selected based on data presented in Tables 23 to 27, Section 3.3.1. The final input parameters for the seven pathogens selected are shown in Table 28.

Table 28: Input parameters of decision trees for the seven pathogens selected

1. Microbial treatment in package with no recontamination?	**Qualitative score**
STEC	No
Salmonella	No
L. monocytogenes	No
Campylobacter	No
Norovirus	No
Cryptosporidium	No
Giardia	No
2. Supports growth?	
STEC	N.A.
Salmonella	N.A.
L. monocytogenes	Yes
Campylobacter	N.A.
Norovirus	N.A.
Cryptosporidium	N.A.
Giardia	N.A.
3. Cooking before consumption?	
STEC	No
Salmonella	No
L. monocytogenes	No
Campylobacter	No
Norovirus	No
Cryptosporidium	No
Giardia	No

N.A.: question not applicable for this pathogen.

3.3.2.2. Risk ranking output

The ranking of the risk metrics provided by the EFSA opinion on public health risks posed by composite foods (2012) is shown in Table 29. The risk was qualified as QPR. Similarly to what was discussed in Section 3.2.1.2 for the first case study, since it is considered that there is a possibility for growth of the pathogen before consumption, and that proper hygienic conditions may not be assumed, the risk should be further qualified as being high for all pathogens.

Table 29: Risk ranking outputs according to the decision tree for the seven pathogens in leafy greens

Product	**Risk**	**Ranking (in both cases)**
STEC	High	1
Salmonella	High	1
L. monocytogenes	High	1
Campylobacter	High	1
Norovirus	High	1
Cryptosporidium	High	1
Giardia	High	1

STEC: Shiga toxin-producing *Escherichia coli*.

3.3.3. EFSA food of non-animal origin risk ranking tool (EFoNAO)

3.3.3.1. Input parameters

The input parameters of the EFoNAO tool are summarised in Table 30. Most of the parameter values were extracted from the EFoNAO opinion (EFSA BIOHAZ Panel, 2013) but since some of the pathogens were not included in the opinion some data were collected from other sources, as indicated in Table 30.

Table 30: Input data for case study. Data are from the EFoNAO opinion unless otherwise stated

	Data	Pathogen						
		Salmonella	***Campylobacter***	**STEC**	***Listeria***	**Norovirus**	***Cryptosporidium***	***Giardia***
Criterion 1	No outbreaks	7	0	0	0	24	0	0
Epidemiological link	No cases	438	0	0	0	657	0	0
	Score	4	1	1	1	4	1	1
Criterion 2	No cases			3 741			6 972[(a)]	167 025[(a)]
Incidence	Multiplier	57.5	Not in opinion	209.6		N.A.	193.5	N.A.
	Total cases	7 117 005	9 000 000	784 166		18 852 364	1 349 034	167 025
	Score	3	3	2	1	4	3	3[(b)]
Criterion 3	DALYs per 1 000 cases	49	40[(c)]	143	2 820[(c)]	2.4	2.9	2.1[(c)]
Public health burden	Score	2	2	3	4	1	1	1
Criterion 4	$IllD_{50}$ (log_{10} CFU)			Not in opinion				
Dose–response	Score	3	3	3	2	3	3	3
Criterion 5	Prevalence	< 1 %	< 1 %	< 1 %	> 1 %	< 1 %	> 1 %	> 1 %
Prevalence of contamination	Score	3	3	3	4	3	4	4
Criterion 6	Percentage consuming	54.2	54.2	54.2	54.2	54.2	54.2	54.2
Consumption	Score	4	4	4	4	4	4	4
Criterion 7	Growth (log_{10}/g) Shelf life (days)							
Growth potential/shelf life	Growth score (G)	3	1	3	3	1	1	1
	Shelf life score (S)	2		2	2			
	Sum of G and S scores	5	1	5	5	1	1	1
	Combined G and S Score	3	1	3	3	1	1	1

$IllD_{50}$: The dose needed to cause illness in 50 % of exposed humans; STEC: Shiga toxin-producing *Escherichia coli*.
(a): TESSy data, 2008.
(b): Assuming same under-reporting as *Cryptosporidium.*
(c): Havelaar et al. (2012).

3.3.3.2. Risk ranking output

The ranking provided by the EFoNAO tool of the selected pathogens is shown in Table 31. The tool ranked *Salmonella* as the highest risk, followed by Norovirus. Then, in order of decreasing risk, two groups resulted; first, STEC and *Listeria*, and then, in the lowest risk group, *Cryptosporidium*, *Giardia* and *Campylobacter*.

Table 31: Scores and ranking of the selected pathogens in leafy greens as estimated by the EFoNAO tool

Criterion	Criterion number	*Salmonella* scores	*Campylobacter* scores	STEC scores	*Listeria* scores	Norovirus scores	*Cryptosporidium* scores	*Giardia* scores
Epi-link	1	4	1	1	1	4	1	1
Incidence	2	3	3	2	1	4	3	3
Public health burden	3	2	2	3	4	1	1	1
Dose–response	4	3	3	3	2	3	3	3
Prevalence of contamination	5	3	3	3	4	3	4	4
Consumption	6	4	4	4	4	4	4	4
Growth potential	7	3	1	3	3	1	1	1
	Sum score	**22**	**17**	**19**	**19**	**20**	**17**	**17**
	Rank	**1**	**4**	**3**	**3**	**2**	**4**	**4**

3.3.4. Risk Ranger

3.3.4.1. Input parameters

The input parameters of Risk Ranger tool were selected based on the dataset presented in Section 3.2. For some parameters, the options provided by Risk Ranger for the values did not match with the dataset. In this case, the option with the closest value to the dataset was selected. The final input parameters related to the risk metric of probability of illness per day per consumer of interest for the seven pathogens are shown in Table 32. Unlike for the *Listeria* example in Section 3.2, in this case the first criterion for hazard severity was not included as it is only related necessary to calculate the risk ranking output of Risk Ranger and not the probability of illness or the total number of illnesses, which were used in the case of leafy greens.

Table 32: Input parameters of Risk Ranger for the seven pathogens

2. How susceptible is the consumer?	**Score**	**Numerical[(a)]**
Salmonella	GENERAL	1
Campylobacter	GENERAL	1
STEC	GENERAL	1
L. monocytogenes	GENERAL	1
Norovirus	GENERAL	1
Cryptosporidium	GENERAL	1
Giardia	GENERAL	1
3. Frequency of contamination		
Salmonella	Common	0.5
Campylobacter	Common	0.5
STEC	Common	0.5
L. monocytogenes	Common	0.5
Norovirus	Common	0.5
Cryptosporidium	Common	0.5
Giardia	Common	0.5
4a. Effect of process		
Salmonella	RELIABLY ELIMINATES hazards	0
Campylobacter	RELIABLY ELIMINATES hazards	0
STEC	RELIABLY ELIMINATES hazards	0
L. monocytogenes	RELIABLY ELIMINATES hazards	0
Norovirus	RELIABLY ELIMINATES hazards	0
Cryptosporidium	RELIABLY ELIMINATES hazards	0
Giardia	RELIABLY ELIMINATES hazards	0

Table 32: Input parameters of Risk Ranger for the seven pathogens (continued)

4b. Effect of preparation for meal		
Salmonella	NO EFFECT on the hazards	1
Campylobacter	NO EFFECT on the hazards	1
STEC	NO EFFECT on the hazards	1
L. monocytogenes	NO EFFECT on the hazards	1
4b. Effect of preparation for meal		
Norovirus	NO EFFECT on the hazards	1
Cryptosporidium	NO EFFECT on the hazards	1
Giardia	NO EFFECT on the hazards	1
5. Is there potential for recontamination?		
Salmonella	OTHER	0.0017
Campylobacter	OTHER	0.0083
STEC	OTHER	0.0054
L. monocytogenes	OTHER	0.0177
Norovirus	OTHER	0.00165
Cryptosporidium	OTHER	0.04
Giardia	OTHER	0.02
6. How effective is the post-processing control system?		
Salmonella	WELL CONTROLLED	1.00
Campylobacter	WELL CONTROLLED	1.00
STEC	CONTROLLED	3.00
L. monocytogenes	CONTROLLED	3.00
Norovirus	WELL CONTROLLED	1.00
Cryptosporidium	WELL CONTROLLED	1.00
Giardia	WELL CONTROLLED	1.00
7. How much increase is required to reach an infectious or toxic dose?		
Salmonella	OTHER	5.05E+01
Campylobacter	OTHER	7.07E+01
STEC	OTHER	3.61E+01
L. monocytogenes	OTHER	1.26E+10
Norovirus	OTHER	3.23E–03
Cryptosporidium	OTHER	8.96E+01
Giardia	OTHER	3.15E+02
8. Frequency of consumption		
Salmonella	Daily	365
Campylobacter	Daily	365
STEC	Daily	365
L. monocytogenes	Daily	365
Norovirus	Daily	365
Cryptosporidium	Daily	365
Giardia	Daily	365

(a): See Section 2.4.2 for description of risk metrics of the tool.

3.3.4.2. Risk ranking output

The ranking of the illness per day per consumer provided by Risk Ranger is shown in Table 33.

Table 33: Risk ranking outputs of Risk Ranger for the seven pathogens in leafy greens

Pathogen	Probability of illness per day per consumer	Ranking
Salmonella	7.71E–07	5
Campylobacter	8.37E–06	3
STEC	8.90E–08	6
L. monocytogenes	2.34E–08	7
Norovirus	7.09E–06	4
Cryptosporidium	3.17E–04	2
Giardia	6.24E–04	1

STEC: Shiga toxin-producing *Escherichia coli*.

3.3.5. microHibro

3.3.5.1. Input parameters

The input parameters used are those described in Section 3.3.1.

3.3.5.2. Risk ranking outputs

The outputs obtained after running the microHibro tool are presented in Table 34.

Table 34: Stochastic risk ranking outputs of microHibro for the mean risk per portion of the pathogens considered

Outcomes	STEC	*Salmonella*	*L. monocytogenes*	*Campylobacter*	Norovirus	*Cryptosporidium*	*Giardia*
Mean probability of illness per day per consumer	6.91E–02	4.23E–03	6.81E–10	7.44E–03	1.36E–01	9.58E–03	3.76E–02
Ranking microHibro	2	6	7	5	1	4	3

STEC: Shiga toxin-producing *Escherichia coli*.

3.3.6. Swift quantitative microbiological risk assessment (sQMRA)

3.3.6.1. Input parameters

The input parameters used are those described in Section 3.2.

3.3.6.2. Risk ranking output

The outputs obtained after running the sQMRA tool are presented in Table 35.

Table 35: Risk ranking outputs of sQMRA for the seven pathogens in leafy greens

Outcomes	STEC	*Salmonella*	*L. monocytogenes*	*Campylobacter*	Norovirus	*Cryptosporidium*	*Giardia*
Mean risk per portion (sQMRA)	1.64E–05	5.30E–06	6.64E–09	5.30E–06	1.70E–04	3.11E–04	4.72E–05
DALYs/ 1 000 cases	143	49	1 450	41	2.4	2.9	2.1
DALYs sQMRA	2.3E–03	2.6E–04	9.6E–06	2.2E–04	4.1E–04	9.0E–04	9.9E–05
Ranks sQMRA	1	4	7	5	3	2	6

DALY: disability-adjusted life years; sQMRA: swift quantitative microbiological risk assessment.

3.3.7. FDA-iRISK

3.3.7.1. Input parameters

The input parameters used are the same as for sQMRA, see Section 3.2.6.1.

3.3.7.2. Risk ranking output

The outputs obtained after running the FDA-iRISK tool are presented in Table 36.

Table 36: Risk ranking outputs of FDA-iRISK for the seven pathogens in leafy greens

Outcomes	STEC	*Salmonella*	*L. monocytogenes*	*Campylobacter*	Norovirus	*Cryptosporidium*	*Giardia*
Mean risk per portion	7.19E–05	7.99E–06	6.60E–09	4.95E–06	1.65E–04	2.94E–04	4.51E–05
DALYs/1 000 cases	143	49	1 450	41	2.4	2.9	2.1
DALYs	0.01	3.9E–04	9.6E–06	2.0E–04	4.0E–04	8.5E–04	9.5E–05
Ranking	1	4	7	5	3	2	6

DALY: disability-adjusted life years; STEC: Shiga toxin-producing *Escherichia coli*.

3.3.8. Burden of Communicable Diseases in Europe (BCoDE)

3.3.8.1. Input parameters

In order to estimate the burden of several selected pathogens transmitted from consumption of leafy greens, we chose to consider the FDA-iRISK outputs on the predicted number of illnesses per serving as the main data source, as this tool provided similar results to the sQMRA tool. For each disease, we distributed the FDA-iRISK output according to the age and sex distribution of the notified cases in the EU, as used in the BCoDE project.

Moreover, we corrected the FDA-iRISK outputs to reflect number of illnesses per 1 million servings, and used this as the main denominator of the BCoDE toolkit (1 million servings = 1 million population). We distributed this population (1 million) across age and sex groups according to European demography.

3.3.8.2. Risk ranking output

The outputs obtained after running the BCoDE tool are presented in Table 37. It is important to note that, at the moment, the BCoDE toolkit is not able to estimate the DALYs of Norovirus as this disease is not part of the BCoDE project. However, it will be possible to create *ad hoc* disease models in a simple building block addition to the toolkit.

Another limitation in this exercise is related to the fact that the BCoDE toolkit only has a model for STEC, not a general *E. coli* model; this might over-estimate the burden of this disease.

Table 37: Risk ranking outputs of BCoDE toolkit for the seven pathogens in leafy greens (the output is very similar to that of FDA-iRISK because similar models are used to calculate the DALYs)

Outcomes	Infection with STEC	Salmonellosis	Listeriosis	Campylo-bacteriosis	Norovirus	Crypto-sporidiosis	Giardiasis
Illnesses per 1 million servings (input from FDA-iRISK)	71.9	7.99	0.0066	49.5	N.A.	294	45.1
DALYs	8.74E+00	3.73E–01	6.52E–02	1.95E+00	N.A.	7.93E–01	1.21E–01
DALYs per 100 000	8.74E–01	3.73E–02	7.57E–03	1.95E–01	N.A.	7.93E–02	1.21E–02
YLD per 100 000	4.88E–01	2.44E–02	6.43E–04	1.75E–01	N.A.	7.89E–02	1.21E–02
YLL per 100 000	3.86E–01	1.29E–02	6.93E–03	1.99E–02	N.A.	4.55E–04	0.00E+00
DALYs per case	1.26E–01	4.66E–02	4.76E+01	3.94E–02	N.A.	2.70E–03	2.68E–03
Ranking (according to DALYs per 100 000)	1	4	6	2	N.A.	3	5

DALY: disability-adjusted life years; STEC: Shiga toxin-producing *Escherichia coli*; YLD: years lived with disability; YLL: years of life lost as a result of premature mortality.

3.3.9. Comparison of the outputs of the risk of multiple pathogens in leafy greens from the different tools

As in the case of the *Listeria* case study, in order to evaluate the performance of the different tools in risk ranking the various pathogens in leafy greens their outputs were compared with a fully quantitative reference risk assessment model which takes into account the main factors affecting the risk and follows the risk assessment paradigm respecting the laws of probability and calculus. The structure of the reference model was the same as used in the *Listeria* case study and the variability of the input parameters was addressed using Monte Carlo simulations using @Risk with 30 000 iterations.

3.3.9.1. Comparison of risk metrics

The overall results of the reference risk assessment model and the different risk ranking tools for the case study of the different pathogens in leafy greens are presented in Table 38.

A comparison in the probability of illness per serving estimated from the reference model and the bottom-up tools FDA-iRISK, sQMRA, microHibro and Risk Ranger is presented in Figure 8.

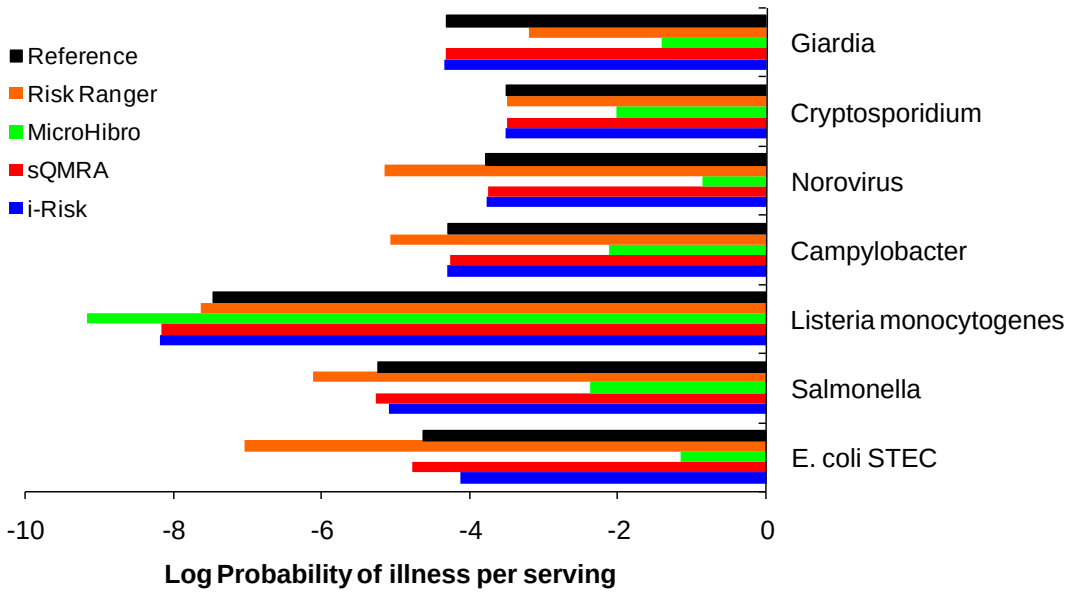


Figure 8: Comparison between the mean probabilities of illness per serving for the different pathogens in leafy greens estimated from the baseline model and the bottom-up tools FDA-iRISK, sQMRA, microHibro and Risk Ranger

In contrast to the *Listeria* case study, in this case FDA-iRISK provided very similar probabilities of illness per serving compared with the reference model for all tested pathogens. This can be attributed to the fact that in this case study growth of all pathogens during storage is limited and the final concentration at the time of consumption does not exceed the maximum population density. As a result, ignoring the latter factor by FDA-iRISK does not affect the output.

The outputs of **sQMRA** were again almost identical to those of the reference model. The significant deviations between the outputs of **microHibro** and the reference model observed in the *Listeria* case study were confirmed in this case of multiple pathogens in leafy greens. The reasons for these deviations remain the calculation problems and the limited number of iterations in Monte Carlo simulation performed with this tool.

As in the *Listeria* case study, **Risk Ranger** provided, in general, lower values for the probability of illness per serving than did the reference model, mainly because Risk Ranger uses the mean values of the input parameters and does not take into account their variability, the simplicity in the dose–response relationship and the limited options for some input parameters.

EFoNAO and decision tree tools provide ordinal or qualitative categorical risk metrics and thus cannot be compared with the other tools with regard to log probability of illness.

The BCoDE estimates of DALYs were slightly different from those derived from FDA-iRISK. Considering that, in this case study, the input of BCoDE was the number of illnesses per 1 million servings estimated by FDA-iRISK, the above differences show the different approach used by BCoDE to estimate the DALYs compared with FDA-iRISK.

3.3.9.2. Comparison of risk rankings

Figure 9 presents a comparison of the rankings of the different pathogens in leafy greens provided by the different tools. The rankings are based on DALYs for FDA-iRISK, sQMRA, microHibro and Risk Ranger, on the sum of scores for EFoNAO and on the categorisation of risk for the decision tree. BCoDE was not included in the comparison since it does not include a disease model for Norovirus;

for the remaining six pathogens the ranking for BCoDE was (1) VTEC/STEC, (2) campylobacteriosis, (3) cryptosporidiosis, (4) salmonellosis, (5) giardiasis and (6) listeriosis.

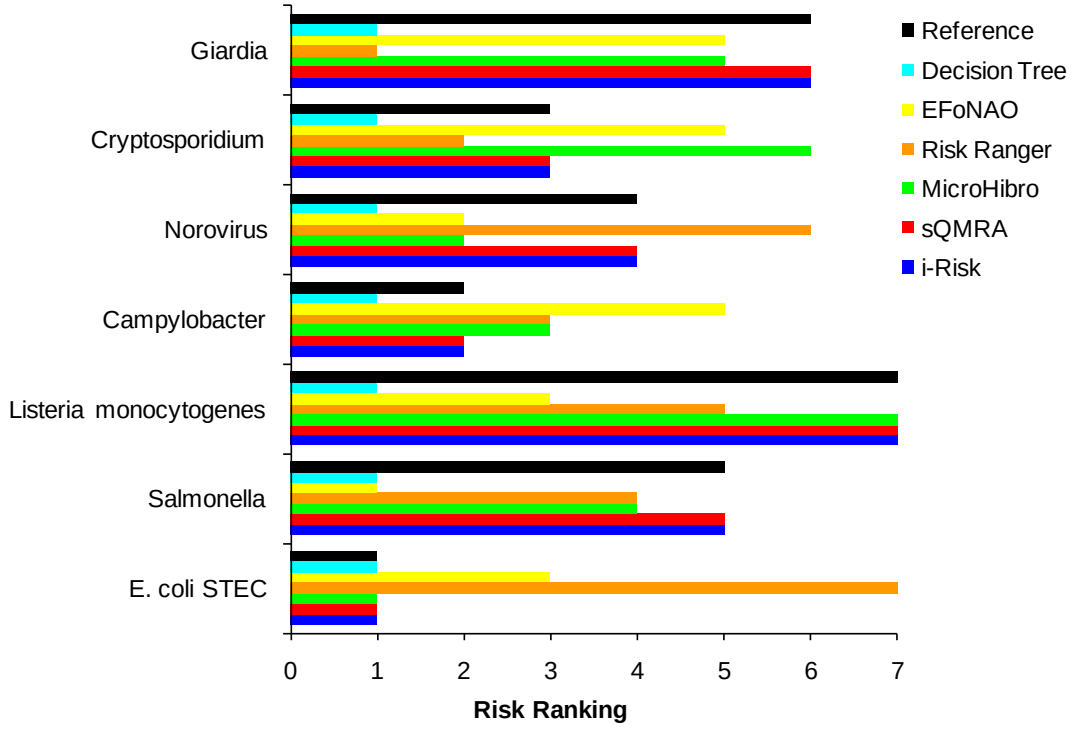


Figure 9: Overall comparison of the rankings of the different pathogens in leafy greens provided by the different tools. The rankings are based on DALYs for FDA-iRISK, sQMRA, microHibro and Risk Ranger, on the sum of scores for EFoNAO and on the categorisation of risk for the decision tree

The reference, the FDA-iRISK and the sQMRA models ranked pathogens identically and in the following order:

STEC > *Campylobacter* > *Cryptosporidium* > Norovirus > *Salmonella* > *Giardia* > *L. monocytogenes*

Excluding Norovirus, the same ranking was also provided by BCoDE, which is probably explained by the fact that FDA-iRISK results were used to feed the BCoDE model.

With microHibro, STEC and *L. monocytogenes* were also ranked first and last, respectively, but the ranking of the rest of the pathogens was completely different. The ranking from Risk Ranger showed the highest deviations from the baseline model compared with the other bottom-up tools. EFoNAO also provided different rankings from the baseline model and showed limited discriminatory capability. The decision tree categorised all of the pathogens as high risk.

The significant dependence of the risk ranking on the selected risk metrics was confirmed in the case study. The following table shows the rankings provided by FDA-iRISK based on the probability of illness per serving and DALYs.

Table 38: Output overview of the tested tools for the risk ranking of *L. monocytogenes* in selected RTE food categories

Tool/risk metric	Pathogen						
	STEC	***Salmonella***	***L. monocytogenes***	***Campylobacter***	**Norovirus**	***Cryptosporidium***	***Giardia***
Baseline model							
Probability of illness/serving	2.35E–05	5.83E–06	3.51E–08	5.22E–05	1.65E–04	3.11E–04	4.81E–05
DALYs[(a)]	3.36E–03	2.86E–04	5.09E–05	2.14E–03	3.96E–04	9.01E–04	1.01E–04
Risk Ranger							
Probability of illness/serving	8.90E–08	7.71E–07	2.34E–08	8.37E–06	7.09E–06	3.17E–04	6.24E–04
DALYs[(a)]	1.27E–05	3.78E–05	3.39E–05	3.43E–04	1.70E–05	9.19E–04	1.31E–03
FDA–iRISK							
Probability of illness/serving	7.19E–05	7.99E–06	6.60E–09	4.95E–05	1.65E–04	2.94E–04	4.51E–05
DALYs[(a)]	1.03E–02	3.92E–04	9.57E–06	2.03E–03	3.96E–04	8.53E–04	9.47E–05
sQMRA							
Probability of illness/serving	1.64E–05	5.30E–06	6.64E–09	5.30E–05	1.70E–04	3.11E–04	4.72E–05
DALYs[(a)]	2.35E–03	2.60E–04	9.63E–06	2.17E–03	4.08E–04	9.02E–04	9.91E–05
EFoNAO							
Total score	19	22	19	17	20	17	17
microHibro							
Probability of illness/serving	6.91E–02	4.23E–03	6.81E–10	7.44E–03	1.36E–01	9.58E–03	3.76E–02
DALYs[(a)]	9.88E+00	2.07E–01	9.87E–07	3.05E–01	3.26E–01	2.78E–02	7.90E–02
Decision tree							
Risk evaluation	High	High	High	High	High	High	High
BCoDE							
DALYs[(a)]	8.74E–03	3.73E–04	6.52E–05	1.95E–03	–	7.93E–04	1.21E–04

DALY: disability-adjusted life years.
(a): DALYs per 1 000 servings.

3.4. Evaluation of tools

The two application examples on *L. monocytogenes* in RTE food categories and on multiple pathogens in leafy greens allowed for a better understanding of the selected risk ranking tools. Based on the above experience, the tools were evaluated according to the following criteria:

- Risk metrics: the ability of a risk ranking tool to provide different risk metrics with meaningful biological or epidemiological interpretation is of great importance. The application examples showed that different metrics can lead to different risk rankings. Thus, it is important to inform the risk managers on which basis metrics (or risk groups) provide a weak scientific basis for risk ranking and may result in misleading outputs.
- Model structure: realistic risk rankings need to be based on models that follow the risk assessment paradigm and respect the laws of probability and calculations.
- Description of input data: the application examples showed that the accuracy in the description of available data as input parameters is an important characteristic of a risk ranking tool.
- Variability and uncertainty: the importance of variability in risk ranking was confirmed by the application examples which showed differences between deterministic and stochastic applications of the tools. The inability of all selected tools to describe uncertainty was also stressed.
- User interface: the experience from the use of the different tools showed that the user interface is important for effective data management, scenario analysis and documentation of the process.

Decision trees use a qualitative approach which permits risk ranking based on descriptive categories of risk (low, moderate, high) with no biological or epidemiological interpretation. The main advantages of the decision trees are that they are able to categorise food–pathogen combinations when limited information is available and are simple to communicate to risk managers. However, because of the structure of the decision trees, it is in practice not possible to include some factors that can significantly affect the final risk. For example, the decision trees used in the application examples of this opinion—selected from previous EFSA opinions—lack a number of significant risk factors, such as the extent of initial prevalence and concentration, extent of growth during storage, the serving size, etc. In addition, arbitrary limits need to be defined in order to split data in arbitrary number of categories for answering the questions of the trees. The above limitations, in combination with the absence of biological or epidemiological interpretation of the risk metric outputs, may result in misleading risk ranking. Furthermore, as confirmed by both application examples, the discriminatory capabilities of decision trees are very limited compared with semi-quantitative and quantitative tools. Uncertainty and variability can be qualitatively described but they are not easily included in the outputs of the decision trees. Although there is no actual user interface, the simple structure of the decision trees allows for easy data management and scenario analysis.

EFoNAO is a semi-quantitative risk ranking tool in an Excel spreadsheet form that uses a mixed bottom-up and top-down approach. Risk ranking with EFoNAO is based on semi-quantitative risk metrics (scores) calculated as the sum of scores of ordinal scoring criteria. The present tool does not take into account factors that can significantly affect the final risk, such as the initial contamination level and the serving size. As a combined bottom-up and top-down approach, the tool provides an evaluation of risk based on certain selected criteria without following the risk assessment paradigm. Advantages of the tool are that the scoring system allows for using qualitative or uncertain input data and that the multi-criterion model is easy to communicate to the risk managers. However, the missing factors that affect the final risk, the ordinal scoring of the criteria, the correlation between some criteria and the lack of a biological or epidemiological interpretation of the risk metric outputs may lead to erroneous risk rankings. EFoNAO does not take into account uncertainty and variability. The Excel spreadsheet requires much manual handling in order to enter, calculate and present results

making data management and scenario analysis complex. However, P^{3}ARRT, which is a tool with the same structure, has a much more advanced user interface.

Risk Ranger is a semi-quantitative risk ranking tool based on a bottom-up approach. It provides meaningful outputs (risk metrics) such as the probability of illness per day per consumer of interest and the total predicted illnesses/annum in population of interest. The main advantage of the tool is that it is simple and easy to use. However, there are a number of weak points in the model's variable and data integration. The serving size, which can be an important factor affecting the final risk, is not included as an input parameter. Serving size can be taken into account only indirectly in the estimation of the increase in the post-processing contamination level that would cause infection or intoxication to the average consumer. The maximum population density of pathogens following growth is also not considered. As a result, the sum of the initial concentration and the growth during retail and domestic storage can be unrealistically high, resulting in over-estimation of risk. Although the model structure and data integration follow, in general, the logic of the standard risk assessment paradigm, there are some weak points. In particular, data integration is simplistic compared with full sQMRA models. For example, a threshold value is assumed for the contamination level that would cause infection or intoxication to the average consumer without taking into account the actual dose–response relationship. For some input parameters the options for their value provided in the risk spreadsheet are limited. In this case, the offered option with the closest value to data must be selected but this can affect the risk ranking. The current version of Risk Ranger is deterministic and does not take into account variability and uncertainty. However, the Excel form of the tool provides flexibility and it could be combined with other software such as @Risk for taking into account variability/uncertainty using Monte Carlo simulation. Guillier et al. (2013) extended Risk Ranger towards a probabilistic version, distinguishing uncertainty and variability. However, this version requires an expert elicitation procedure in which the expert is asked for two quantiles to assess variability as well as given quantiles to incorporate an uncertainty level. Data management and scenario analysis with Risk Ranger is complex. Each scenario (pathogen–product pair and/or differences in input parameters) requires a different file to be stored which complicates quality assurance evaluation and comparison of different scenarios.

FDA-iRISK is a quantitative, bottom-up risk assessment tool providing meaningful risk metrics such as the probability of illness per serving, the total annual number of illnesses and DALYs, which can be used for risk ranking. The main weak point of the current version of FDA-iRISK is that it does not take into account the maximum population density of pathogen's growth, which may result in an unrealistically high concentration of the pathogens at the time of consumption and over-estimation of risk. Apart from the above weakness, the tool takes into account the main factors affecting the risk and follows the risk assessment paradigm respecting the laws of probability and calculus. The user can run the tool in both a deterministic and a stochastic way. For the stochastic applications, various probability distributions are available for describing input data (fixed, normal, beta-PERT, uniform, triangular, uniform and empirical cumulative distribution). The tool accepts only input data describing the increase or decrease of concentration and prevalence, while specific growth or inactivation models have to be run outside the tool. An advantage of FDA-iRISK is that the number of iterations is automatically selected based on simulation convergence criteria and not settled before by the user. All probability distributions are assumed to describe variability since the current version does not include uncertainty. The FDA-iRISK tool has the more advanced user interface among the tested tools in this opinion. It is capable of modelling different steps in the food chain from farm to fork providing flexibility in choosing different scenarios combining hazards, consumption patterns and processing stages. In addition, each model run can be saved and shared online with other users, allowing effective quality assurance evaluation and comparison of different scenarios.

sQMRA is a quantitative, bottom-up risk assessment tool in an Excel spreadsheet form that can be used for risk ranking based on various meaningful risk metrics including probability of illness per serving, total annual number of illnesses, DALYs and cost-of-illness. sQMRA takes into account all the factors affecting the risk and follows the risk assessment paradigm respecting the laws of probability and calculus. The tool can provide both deterministic and stochastic outputs for risk

ranking using single values or distributions for the input parameters, respectively. However, in the latter case only a limited number of probability distributions is available for describing these data. This limitation may lead to erroneous ranking outputs when input data are not in a form that can be described by an available probability distribution. An advantage of the tool is that growth of the pathogens during storage can be estimated within the tool using the appropriate parameters in a secondary cardinal model. In the stochastic application, the number of iterations in the Monte Carlo simulation procedure has to be settled in advance without taking into account simulation convergence criteria. This may result in differences in the outputs for different number of iterations and between different simulations. All probability distributions in sQMRA are assumed to describe variability since the current version does not include uncertainty. The Excel spreadsheet form of the tool provides an informative summary of input data and allows for adequate checks on input validity. However, a weak point of the tool is that the spreadsheet form makes file management very complex with each scenario (pathogen–product pair and/or differences in input parameters) requiring a different file to be stored which complicates quality assurance and comparison of different scenarios.

microHibro was initially developed as a microbial growth prediction tool, but with recent developments the model can be used for quantitative risk assessment and risk ranking. In its current form, the tool can estimate only the probability of illness and the number of illnesses. It takes into account all the factors affecting the final risk following the risk assessment paradigm and respects the laws of probability and calculus. The user can run the tool only in a stochastic way since the deterministic application cannot take into account the prevalence of the pathogens. Various probability distributions are available for describing input data (normal, gamma, uniform, exponential, triangular, Poisson). An advantage of microHibro is that growth or inactivation of the pathogens can be estimated within the tool using the appropriate growth model. In the stochastic application, the number of iterations in the Monte Carlo simulation procedure has to be set in advance, without taking into account simulation convergence criteria. In the current version of the tool, the Monte Carlo process is very slow and may result in differences in the outputs for different number of iterations and between different simulations. All probability distributions are assumed to describe variability since the current version does not include uncertainty. The microHibro has an advanced user interface and the user can design any step in the food chain from farm to fork. The advanced interface allows for effective data management and analysis of different scenarios combining hazards, consumption patterns and processing stages. Furthermore, both risk assessment and growth/inactivation models can be saved and shared online with other users. However, the development for a risk assessment application is in progress and there is a need for further improvements in the calculations and the presentation of the results.

BCoDE is a full top-down risk ranking tool that provides meaningful outputs such as DALYs, DALYs per case, DALYs per 100 000, YLD and YLL per 100 000. Risk ranking with BCoDE is based on a limited number of input parameters, namely the age group- and sex-specific number of cases, which reduces complexity of the tool. Flexibility is ensured by the possibility of changing all other parameters, such as population data (as in the listeriosis case study of this opinion), life expectancy and all parameters of the disease models (disability weights, transition probabilities and durations). Variability and uncertainty of all variables (number of cases, disease model variables, population data) are taken into account using Monte Carlo simulations (up to three inputs are possible for each variable) and outputs include mean, median, 2.5^{th} and 97.5^{th} percentiles. BCoDE has an advanced, user-friendly and intuitive interface that allows effective data management and scenario analysis while outputs are presented in communication-friendly visualisations such as tables, bubble charts and bar charts. However, BCoDE is able to estimate DALYs only from incidence data and does not take into account transmission pathways: translation of source attribution to incidence must be performed beforehand by the user. For a more specific application in a food safety context, incidence estimates are needed at a higher level of resolution, i.e. for specific food or group of foods within population subgroups. These estimates can be provided by attribution models which are not currently included in the tool. Alternatively, as we show in the application case study of this opinion, BCoDE can be used in combination with a bottom-up risk ranking tool. In this case, the number of illnesses for a specific

food or food category estimated with a bottom-up approach can be used as an input in BCoDE for a more effective estimation of DALYs.

4. Comparison of general risk ranking approaches: stochastic, deterministic and ordinal scoring

The risk ranking tools evaluated in this opinion are based on different approaches including qualitative, semi-quantitative with ordinal scoring, quantitative deterministic and quantitative stochastic. Because of the additional differences found between the tools other than the approach (see Section 2), their comparison presented in Section 3 cannot provide adequate information about the performance of the above approaches in risk ranking. In addition, there are no studies available in the literature providing a comparative evaluation of these approaches. The objective of this section was to systematically compare stochastic, deterministic and ordinal scoring approaches in risk ranking.

4.1. Methodology of comparison

For the purpose of comparison of the different approaches, a generic stochastic risk assessment model from retail to consumption was defined. A probability distribution was selected for each variable of the model for the description of variability. In each parameter of the above distributions, a range of values was assigned to cover different food hazards characteristics. By randomly selecting a value from the above ranges, a dataset of the model input parameters for food–pathogen combinations can be generated. Several hundreds of datasets representing a corresponding number of food–pathogen combinations were generated and the risk of each combination was estimated using stochastic, deterministic and ordinal scoring approaches. In the stochastic approach, the variables of the model were described with probability distributions and the final risk was estimated using Monte Carlo simulation. In the deterministic approach, the variables of the model were described with single values using different statistical measures (i.e. arithmetic mean, median, 75^{th} percentiles and 90^{th} percentiles) for comparison. For the ordinal scoring approach, a score was assigned to the variables of the model based on their categorisation on a continuous scale. The overall score was obtained by summing the scores assigned to each variable.

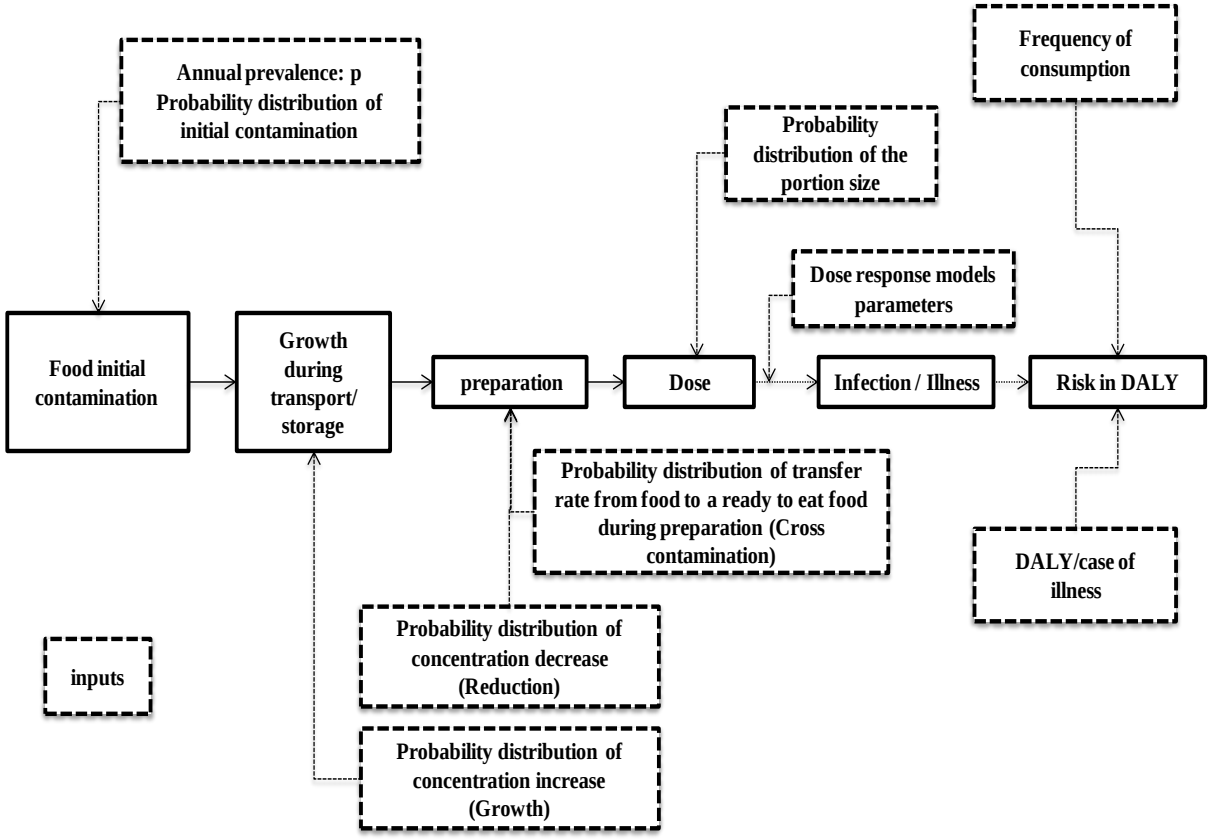
The ranking of the food–pathogen combinations derived from the different approaches were compared both graphically and using appropriate statistical measures. Assuming that the stochastic approach provides the most realistic outputs since it takes into account the variability of the risk determinants, the deterministic and ordinal scoring approaches were evaluated in relation to the stochastic one.

4.1.1. Generic risk assessment framework

There are many ways in which risk, and the individual factors of risk, have been defined and evaluated. When reliable quantitative data are available, quantitative multiplicative mathematical model may be used to estimate risk. From retail to consumption, the changes in concentration of pathogens in the food are described using the available predictive microbiology models in combination with the probability distributions of the temperatures of the food during transport and storage of the food product. During storage and preparation, microorganisms present in one food product can be transferred to a RTE food (cross-contamination). The range of possible transfer rate values depend on the food characteristics and food handling by the consumer. Hoelzer et al. (2012) synthesised available data and derived probability distributions and mathematical models of bacterial transfers between food and environmental surfaces and vice versa. When food products are cooked, the survival of microorganisms is described thanks to the available predictive models in combination with the probability distribution of temperatures and durations of cooking. To assess the concentration at time of consumption from this series of distributions, Monte Carlo simulations are performed. The exposure dose, number of pathogen cells in an ingested serving, is determined from the final concentration of pathogens reached after the accumulated growth or survival calculated at each step and the quantity of food product. A dose–response model is then applied to calculate the probability of infection/illness from that number of pathogen cells in a single serving. The total number of illnesses in a population can be calculated by multiplying the mean probability of illness per serving by the total

number of servings consumed by the population. Finally, the public health impact can be estimated by translated the total number of illnesses in DALYs.

The structure of the generic risk assessment framework is represented in the Figure 10. Models following this generic framework predict, from the initial contamination level (at the time the products leave the retail stores), the potential amount of microbial hazard to be consumed under a wide range of situations. To cover all relevant situations, a stochastic modelling approach is used, where variable are included and described by probability distributions of possible values rather than a single estimate (Table 39). At this stage only variability is included and the parameters of the probability distributions were assumed as perfectly known.



DALY: disability-adjusted life years.

Figure 10: Risk assessment framework and inputs

The generic risk assessment framework assumes the following:

- Initial contamination at retail level (H_0): characterised by three parameters, p (prevalence, proportion of contaminated food product units), μ_0 (mean of the concentration in food in $\log_{10}$ CFU/g or ml) and σ_0 (the standard deviation of concentration in food in $\log_{10}$ CFU/g or mL). The concentration is assumed to be log-normally distributed.
- Change in concentration during transport to home and storage: (ΣG) characterised by a log-gamma distribution with parameters μ_G and σ_G and derived from a predictive model with maximum population density.
- Cross-contamination during preparation (ΣC): characterised by a log-normal distribution with parameters μ_c and σ_c. It is assumed that a fraction of the microbial hazards present in the handled food product unit is transferred to a RTE food and all the transferred amount of the microorganism will be ingested by the consumer.

- Change in concentration during preparation (e.g. cooking) (ΣR): assumed to be log-normally distributed with parameters μ_R and σ_R.
- Portion size (S): characterised by a gamma distribution with a mean and standard deviation noted respectively as μ_S and σ_S.
- Dose–response model: an exponential dose–response model with fixed parameter r is used.
- Consequence function: average DALYs per case is used.
- Population at risk: we used the average number of eating occasions per year per person.

The mathematical equations and their combinations are presented in Table 39.

Table 39: Generic risk assessment framework description

Variables	Unit	Distribution/formula	Input parameters
Initial concentration (H_0)	Log_{10} CFU/g	Normal	μ_0 and σ_0
Portion size	g	Gamma (α,β)	$\mu_s = \alpha\beta$ $\sigma_s = \beta\sqrt{\alpha}$
Expected CFU per portion (E_0)	CFU/portion	$E_0 = S \times 10^{H_O}$	
Increase during storage (G)[(a)]	Log_{10}	Gamma (a,b)	$\mu_G = \alpha\beta$ $\sigma_G = \beta\sqrt{\alpha}$
Expected CFU per portion end of storage (E_S)	CFU/portion	$E_S = E_0 \times 10^G$	
CFU per portion end of storage (X_S)	CFU/portion	Poisson (E_s)	
Log_{10} probability of transfer to RTE (C)	Log_{10}	Normal	μ_c and σ_c
CFU transferred per portion (D_1)	CFU/portion	Binomial (X_s, 10^C)	
CFU remaining per portion (X_{nc})	CFU/portion	$X_{nc} = X_s - D_1$	
Log_{10} probability of survival during cooking	Log_{10}	Normal	μ_R and σ_R
CFU surviving cooking (D2)	CFU/portion	Binomial (X_s, 10^R)	
Probability of infection (PInf)		$PInf = 1 - (1 - r)^{(D1 + D2)}$	r
Probability of illness (PIll)		PIll = PInf × P(Ill | infection)	P(Ill | infection)
Average probability of illness (APIll) per contaminated serving		Arithmetic mean of probability of illness (Monte Carlo simulation, 50 000 iterations)	
Annual probability of illness (API)		API = P × APIll × FR	FR: average number of eating occasion per year per person P: prevalence
Annual DALYs per 1E6 consumers		ADALY=API × DALY × 1E6 consumers	DALY per case

DALY: disability-adjusted life years.
(a): Based on relevant predictive modelling.

4.1.2. Generation of datasets for food–pathogen combinations

In order to describe the differences in the various food–pathogen combinations, a range of values was given to each input parameter of the variables in the generic framework presented in Table 39. The ranges of values of the parameters are shown in Table 40. By randomly selecting a value from the above ranges, a dataset of the input parameters for each food–pathogen combination was generated. Initially, 700 datasets representing a corresponding number of food–hazards combinations were generated. Further, the risk for these 700 combinations was assessed using a stochastic modelling approach, which followed the generic framework presented in Figure 10. The results showed that, for some food–pathogen combinations, the estimated risk was unrealistically high or low. In order to make the example more realistic, 392 out of the total 700 combinations were selected and included in

the analysis, for which the final risk of illness per year per person was between 10^{-12} and 0.8, and the corresponding DALYs were lower than 300 per 1 000 000 persons per year. This DALY reference was obtained using the reported salmonellosis incidence rate, 22/100 000 cases (EFSA and ECDC, 2014) and combining it with an average DALY value of 49/1 000 and an underreporting factor of 30.

Table 40: Range of parameters to generate input data for food–pathogen combinations in the stochastic model used

Variables	Unit	Parameters	Ranges of the parameters values
Initial concentration (H_0)	Log_{10} CFU/g	μ_0 σ_0	–3 to 3 0.1 to 1.5
Prevalence		P	10^{-4} to 1
Portion size	g	μ_s σ_s	10 to 500 0.1 to 1
Increase during storage (G)	Log_{10}	μ_g σ_g	0.3 to 3 0.1 to 1.5
Log_{10} probability of transfer to RTE (C)	Log_{10}	μ_c σ_c	–5 to –2 0.1 to 1.5
Log_{10} probability of survival during cooking	Log_{10}	μ_R σ_R	–6 to –3 0.1 to 1.5 If RTE product (50 % of the simulated scenario R = 0)
Probability of infection (PInf per CFU)		r	– 10 to – 2
Probability of illness (PIll)		PIll = PInf × P(Ill | infection)	1
Average number of eating occasions per year per person		FR	1 to 365
DALY per case	Year (log_{10})	DALY	–3 to 1

DALY: disability-adjusted life years.

4.1.3. Risk ranking comparison

The stochastic approach was considered as the reference risk ranking approach and the deterministic and ordinal scoring approaches were evaluated by comparing their rankings with that of the stochastic approach.

Both the Spearman rank correlation coefficient and Kendall's tau correlation coefficient were used measures for comparing the rankings. Spearman rank correlation coefficient was first proposed as a non-parametric rank statistics to measure the strength of association between two variables. It is defined as:

$$r = 1 - \frac{6}{N^3 - N}\sum_{i=1}^{N} d_i^2$$

where d_i is the difference between the ranks of items i and N is the number of ranked items. Two rankings are identical when the coefficient is 1, and in inverse order when the coefficient is –1.

The Kendall's tau rank correlation coefficient is defined as:

$$\tau = 2\,p - 1$$

$$p = \frac{C}{N(N-1)/2}$$

where C is the number of concordant pairs (pairs that are ranked in the same order in both rankings) and N is the number of ranked items. Note that if two rankings are identical (p = 1), then Kendall's tau value is 1, whereas, if the two rankings totally disagree (p = 0), then Kendall's tau value is –1, and if the two rankings are independent (p = 1/2), then Kendall's tau value is 0.

Kendall's tau can be used to find which method is better relative to "gold standard". The higher Kendall's tau value that measures the correlation between the output ranking of a method and the gold standard, the better the method is concluded to be. Pairs of ranking whose Kendall's tau value are equal or higher than 0.9 can be considered "effectively equivalent".

Although Kendall's tau is considered as a useful measure for comparing two rankings, there is an important problem with this statistic. Kendall's tau equally penalises errors that occur at any part of the ranked list. Therefore, Kendall's tau does not distinguish between the errors that occur towards the top of the list from the errors towards the bottom of the list. Since the food–pathogen pairs that are placed at the top of the list are more important than those towards the bottom, there is a need to find a measure that assigns more weight to the errors made towards the top of ranking than to the errors towards the bottom. Yilmaz et al. (2008) proposed a new rank correlation coefficient based on the principle of average precision.

The average precision (AP) rank correlation coefficient is calculated as following:

$$\tau_{AP} = 2\,p' - 1$$

$$p' = \frac{C}{N-1}\sum_{1=2}^{N}\frac{C(i)}{(i-1)}$$

where C(i) is the number of items above the rank i and correctly ranked with respect to the item at rank i. Note that p′ is very similar to the p upon which Kendall's tau is based; the only difference is that, instead of comparing an item with any other ranked item, it is compared only with items above. The values of p′ fall between 0 and 1, where 1 means that all items ranked by a method are ranked in the same order as the items ranked by the reference method and 0 means that all items ranked above another item are ranked incorrectly according to the reference method. The average precision rank correlation coefficient values will fall between –1 and +1 and interpreted in the same manner as Kendall's tau.

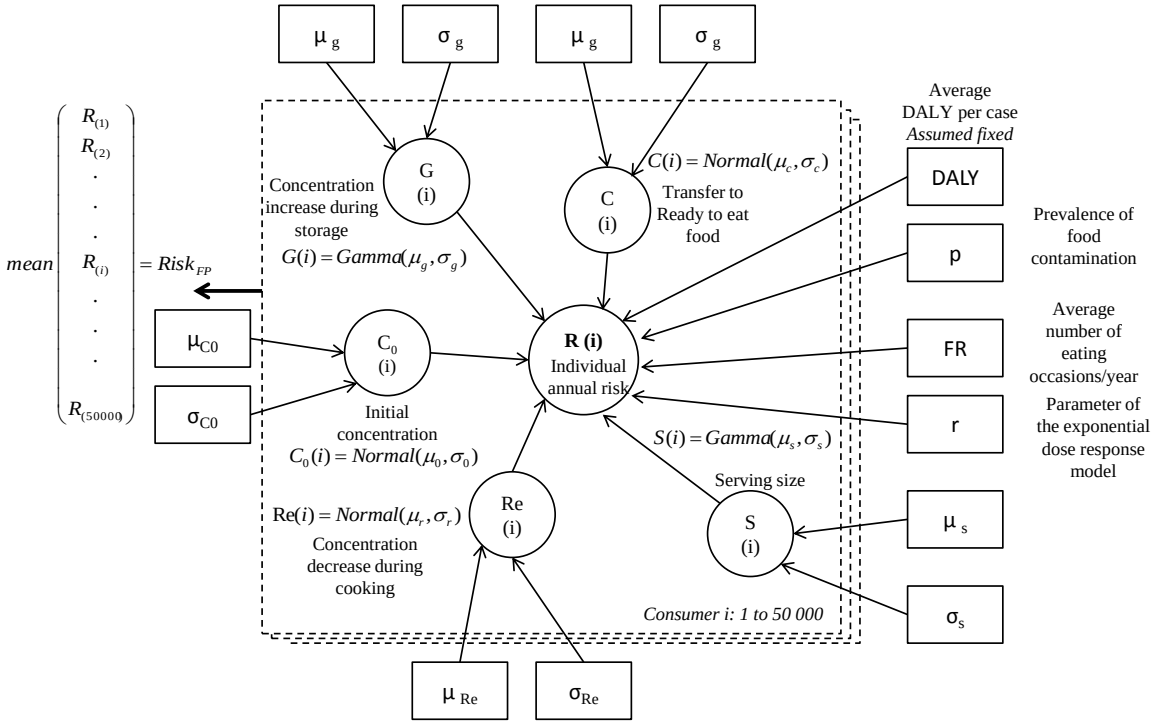
In principle, when the ranking errors are uniformly distributed over the list, Kendall's tau and the average precision rank correlation coefficient are equivalent. When there are more errors towards the top of the list, then Kendall's tau is always greater than the average precision rank correlation coefficient ($\tau > \tau_{AP}$), and, when there are fewer errors towards the top of the list, $\tau < \tau_{AP}$.

4.2. Results

4.2.1. Stochastic risk ranking approach: the reference approach

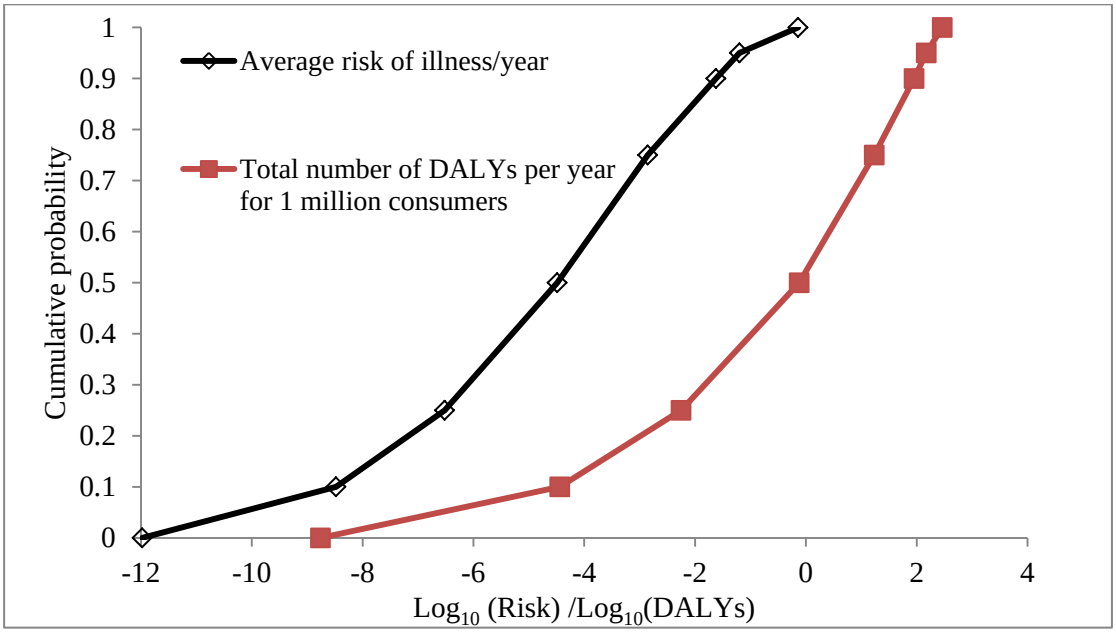
The structure of the stochastic application of the generic risk assessment framework, which was used as the reference approach, is presented in Figure 11. In the stochastic approach, each variable of the model was described with a probability distribution (Table 39) and the risk of each food–pathogen combination ($Risk_{FP}$) expressed as average probability of illness per year and total number of DALYs per year for 1 million consumers was estimated using Monte Carlo simulation with 10 000 iterations.

Figure 12 presents the cumulative probability of illness per year for the 392 food–pathogen combinations. The relationship between the ranking of the 392 food–pathogen combinations and their risk expressed in total number of DALYs per year for 1 million consumers is presented in Figure 12. Table 41 shows the statistics of the average probability of illness per year and total number of DALYs per year for 1 million consumers estimated with the stochastic approach.



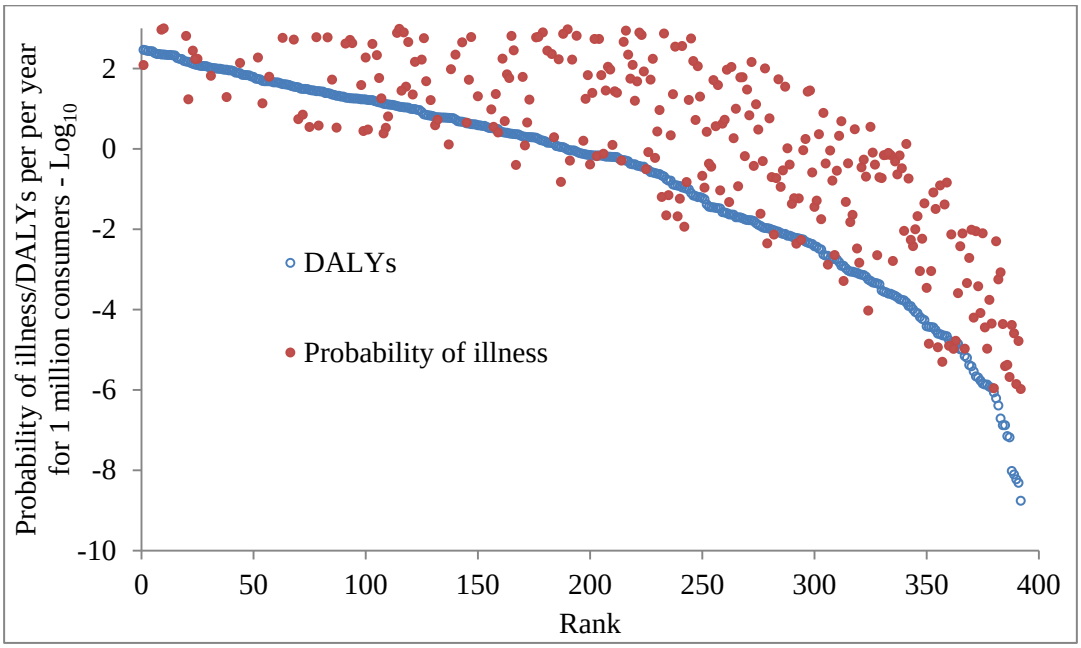
DALY: disability-adjusted life years.

Figure 11: Structure of the reference stochastic approach applied to the generated datasets of the food–pathogen combinations. For each food–pathogen combination a single risk measure is derived: $Risk_{FP}$ (expected total number of DALY per year for 1 million consumers). Circles represent random variables and rectangles fixed values.



DALY: disability-adjusted life years.

Figure 12: Cumulative probability of illness per year for the 392 food–pathogen combinations



DALY: disability-adjusted life years.

Figure 13: Relation between the ranking of the 392 food–pathogen combinations and their risk expressed in total number of DALY per year for 1 million consumers

Table 41: Statistics of the assessed the assessed risk 392 food–pathogen combinations

Statistics	Average risk of illness/year	Total number of DALYs per year for 1 million consumers
Minimum	1.05E–12	1.74E–09
Maximum	7.24E–01	2.90E+02
10th percentile	3.31E–09	3.61E–05
25th percentile	3.03E–07	5.61E–03
50th percentile	3.26E–05	7.58E–01
75th percentile	1.40E–03	1.71E+01
90th percentile	2.39E–02	8.97E+01
95th percentile	6.32E–02	1.49E+02

DALY: disability-adjusted life years.

In order to find out which parameters are influencing the obtained ranking the most, we calculated as a first approach the Kendall tau b correlation coefficients between the model outputs and the input parameters (Figure 14 and Table 42). The model seems to be more sensitive to the dose–response model, growth potential, initial concentration, reduction during cooking and DALY parameters. Moreover, particularly noteworthy is the fact that the rank of a particular food–pathogen pair is influenced by the mean and standard deviation of the distribution describing the variability. This shows clearly the expected bias on the risk estimates if variability is ignored.

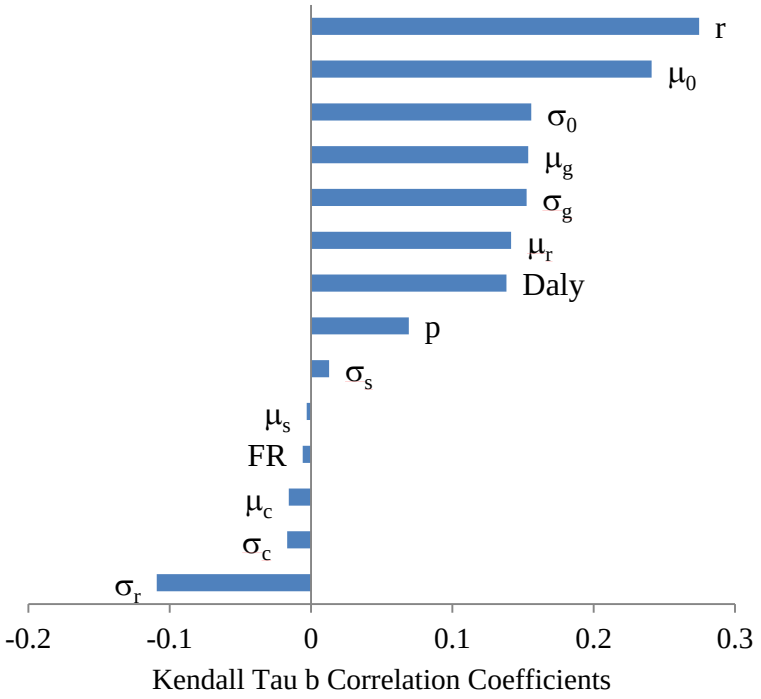


Figure 14: Kendall tau b correlation coefficients between the assessed risk (stochastic approach) and the input parameters, σ_r: potential reduction (standard deviation), μ_c: cross-contamination (mean), FR: frequency of consumption, σ_c: cross-contamination (standard deviation), σ_s: portion size (standard deviation), μ_s: portion size (mean), p: prevalence, μ_r: potential reduction (mean), σ_0: initial concentration (standard deviation), Daly: DALY (disability-adjusted life years), μ_g: growth potential (mean), σ_g: growth potential (standard deviation), μ_0: initial concentration (mean), r: dose–response model parameter

Table 42: Kendall tau b correlation coefficients between the assessed risk (stochastic approach) and the input parameters

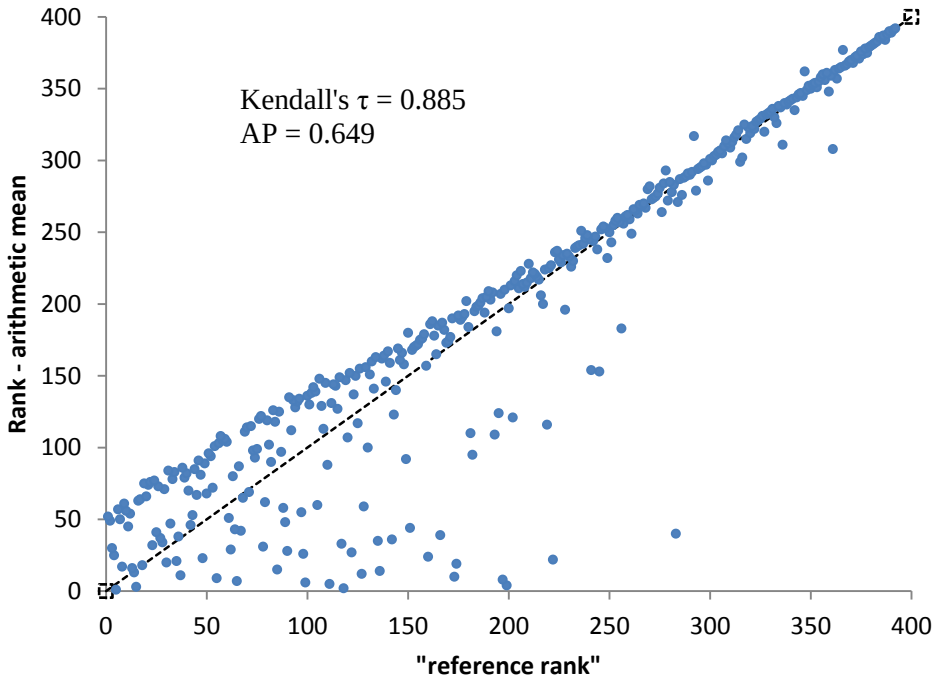
Parameters	Notation	Kendall tau b correlation coefficients
Potential reduction (SD)	σ_r	–0.109
Cross-contamination (SD)	s_c	–0.017
Cross-contamination (mean)	m_c	–0.016
Frequency of consumption	FR	–0.006
Portion size (mean)	m_s	–0.003
Portion size (SD)	s_s	0.013
Prevalence	p	0.069
DALY	DALY	0.138
Potential reduction (mean)	m_r	0.141
Growth potential (SD)	s_g	0.152
Growth potential (mean)	m_g	0.154
Initial concentration (SD)	s_0	0.156
Initial concentration (mean)	m_0	0.241
Dose–response parameter	r	0.274

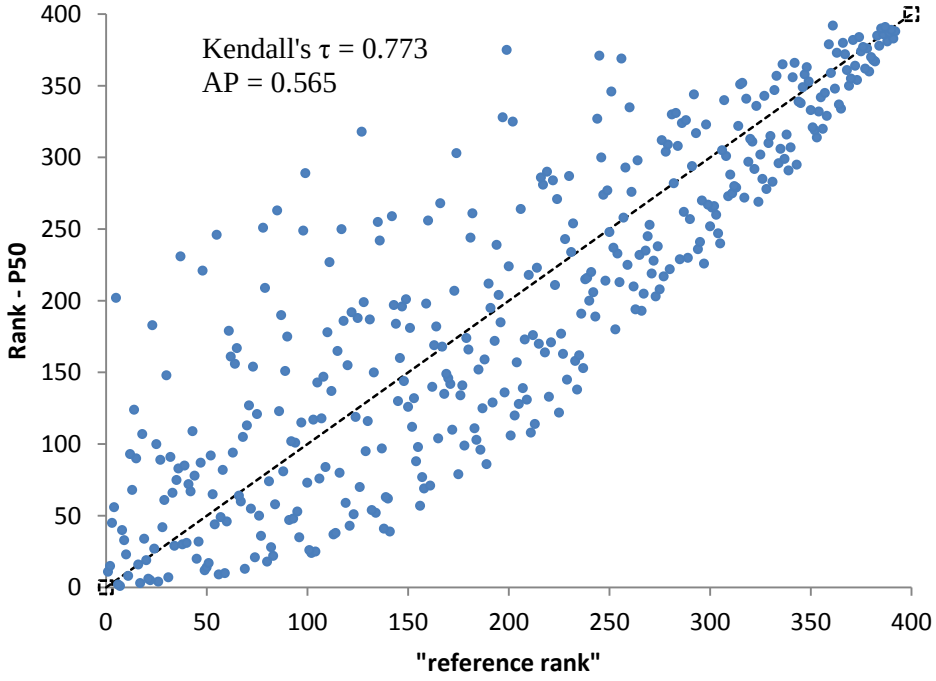
SD: standard deviation; DALY: disability-adjusted life years.

4.2.2. Evaluation of the deterministic risk ranking approach

Most of the current risk rankings of the food–pathogen pairs have been carried out using methodologies that do not allow the inclusion of the variability and uncertainty inherent to food, pathogen and consumer variables. As a consequence, such methodologies can provide only a very limited (and potentially biased) assessment of the relative risk associated with the food–pathogen pairs.

A model following the generic risk assessment framework was used in a deterministic way to check if the deterministic approach leads to equivalent ranking as the stochastic approach. Instead of using probability distributions, as in the stochastic approach for the initial concentration, growth potential, cross-contamination probability, reduction during cooking and serving size, single values are used. In order to investigate the effect of using different statistical measures of the probability distributions describing the model's variables, the arithmetic mean, median, 75^{th} percentiles or 90^{th} percentiles were evaluated. Further, the ranking obtained with the stochastic model named as "reference rank" was compared with the four rankings derived from the deterministic approach named Rank(arithmetic mean), Rank(P50), Rank(P75) and Rank(P90) in which the arithmetic mean, median, 75^{th} percentiles and 90^{th} percentiles were used, respectively, as inputs. The comparison was performed both graphically and using the Kendall's tau and AP indexes. Figure 15 shows the discrepancy between the deterministic and stochastic risk rankings.


Kendall's τ = 0.885
AP = 0.649
Rank - arithmetic mean
"reference rank"
0
50
100
150
200
250
300
350
400


Kendall's τ = 0.773
AP = 0.565
Rank - P50
"reference rank"
0
50
100
150
200
250
300
350
400

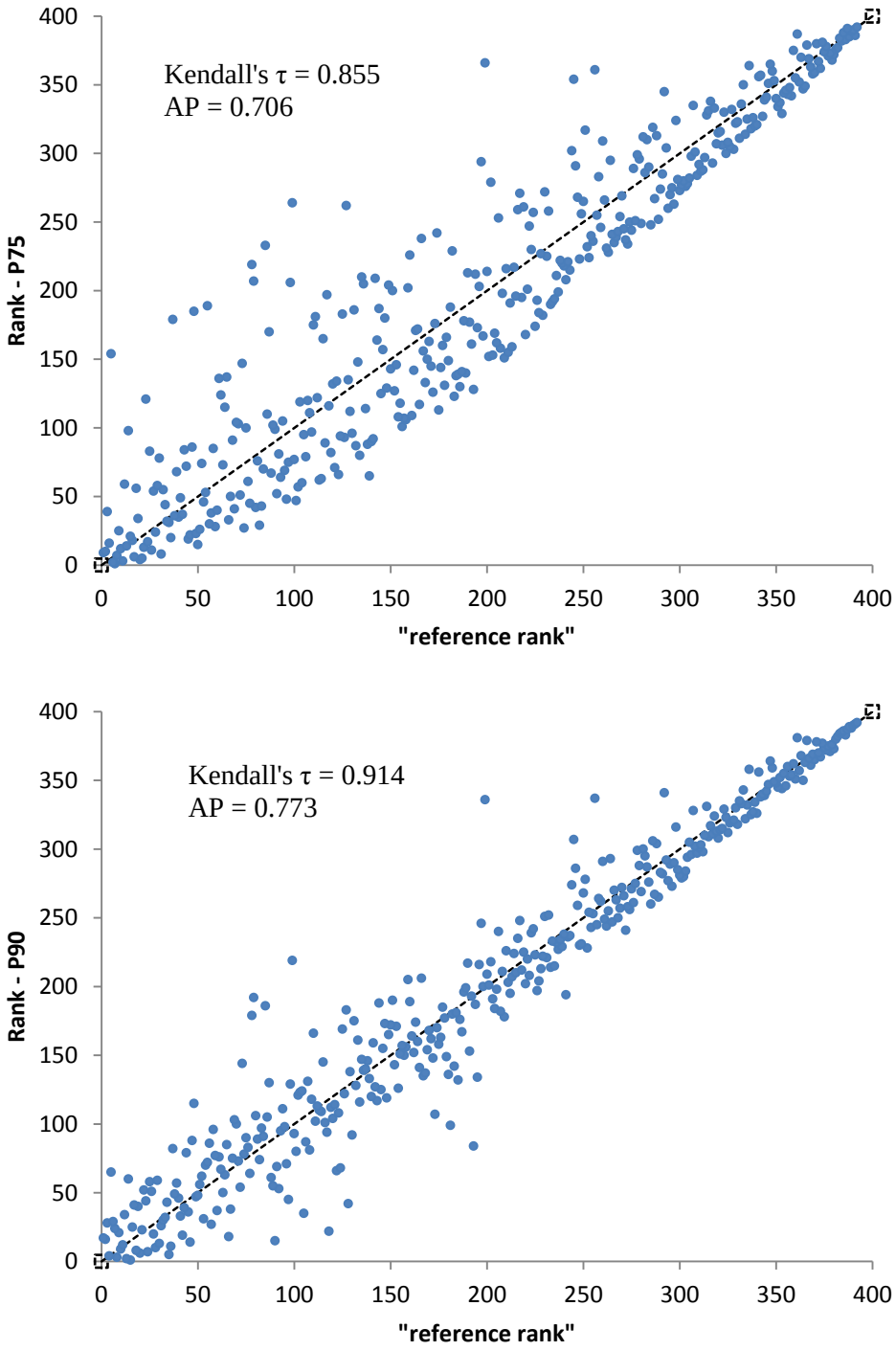


Figure 15: Comparison between the ranking obtained with the stochastic model ("reference rank") and the four rankings derived from the deterministic approach named as Rank(arithmetic mean), Rank(P50), Rank (P75) and Rank(P90) in which the arithmetic mean, median, 75th percentiles and 90th percentiles were used, respectively, as inputs

According to the Kendall's tau coefficient, the highest differences between deterministic and stochastic are observed when the model was run with median as input (t = 0.773, Figure 15). The Kendall's tau values that measure the correlation between the output rankings provided by the deterministic approaches and stochastic approach were 0.914, 0.885, 0.855 and 0.773 for rankings using P90, arithmetic mean, P75 and P50, respectively. One can conclude that, in general, the use of P90 provides the closest ranking to the reference approach. As shown in the graphs, the difference

between the deterministic and stochastic approaches varies between different food–pathogen combinations, i.e. some food–pathogen combinations are positioned close to the diagonal line, indicating that the ranking between deterministic and stochastic approaches are very close. As a consequence, the performance of the different statistical point estimates (means or percentiles) used in the deterministic approach will depend on the specific food–pathogen combination involved in the ranking and their position in the risk range. All the calculated average precision rank correlation coefficient (AP, Figure 15) are lower than the Kendall's tau coefficient, showing that all the deterministic approaches have more errors towards the top of the list when compared with the errors towards the bottom of the list. Indeed, even in the case of the Rank(P90), which showed the best performance, some miss ranking can be obtained, i.e. for some combinations ranked close to 1 with the stochastic model (highest risk combination) the deterministic approach may rank such combination at lower risk as a difference of –50 in the rank can be obtained (under-estimation).

4.2.3. Evaluation of the semi-quantitative risk ranking approach with ordinal scoring

Semi-quantitative risk assessment models with ordinal scoring provide an intermediary level between the textual evaluation of qualitative risk assessment and the numerical evaluation of quantitative risk assessment, by evaluating risks with a score. The ordinal scoring approach does not require the same mathematical skills as for quantitative assessments and can be applied with less precise data. The system for assignment of a category for a food–pathogen combination used in this example uses nine criteria: initial concentration, prevalence, portion size, number of eating occasions, increase during storage, transfer to RTE during food handling, reduction during cooking, dose–response model and DALYs per case. For each variable, quantitative inputs on a continuous scale were assigned to a limited number of categories. The categories were in general defined using a logarithmic scale, as shown in Table 43. The ordinal scores were defined in a linear (arbitrary) scale from 1 to 5 or using a logarithmic transformation based on the formula:

$$\frac{\ln(x')}{(1-\ln(x'))}$$

where $x' = x/(x_{max} - x_{min})$ and × = bin limit (Havelaar et al., 2010). The overall score was obtained by adding the scores assigned for each criterion.

Table 43: Categories and scores defined in the ordinal scoring approach

Inputs	Bins(x)	Ordinal score		Inputs	Bins(x)	Ordinal score	
		Linear	Log-scaled			Linear	Log-scaled
Initial concentration (H_0) in CFU/g	1.0E–03	1	0.000	Prevalence	1.0E–04	1	0.000
	1.0E–02	2	0.200		1.0E–03	2	0.250
	1.0E–01	3	0.400		1.0E–02	3	0.500
	1.0E+00	4	0.600		1.0E–01	4	0.750
	1.0E+01	5	0.800		3.0E–01	5	0.869
Portion size in grams	1.0E+01	1	0.000	Average number of eating occasions per year per person	1.0E+00	1	0.000
	3.0E+01	2	0.239		1.2E+01	2	0.421
	9.0E+01	3	0.477		5.2E+01	3	0.670
	2.7E+02	4	0.716		1.0E+02	4	0.787
	8.1E+02	5	0.954		2.1E+02	5	0.905
Increase during storage (G)	1.0E+00	1	0.000	Probability of transfer to RTE (C)	1.0E–05	1	0.000
	1.0E+01	2	0.200		1.0E–04	2	0.200
	1.0E+02	3	0.400		1.0E–03	3	0.400
	1.0E+03	4	0.600		1.0E–02	4	0.600
	1.0E+04	5	0.800		1.0E–01	5	0.800

Table 43: Categories and scores defined in the ordinal scoring approach (continued)

Inputs	Bins(x)	Ordinal score		Inputs	Bins(x)	Ordinal score	
		Linear	Log-scaled			Linear	Log-scaled
Reduction during cooking	1.0E+00	1	0.000	Probability of infection (PInf per CFU)	1.0E–12	1	0.000
	1.0E+01	2	–0.125		1.0E–06	2	0.500
	1.0E+02	3	–0.250		1.0E–04	3	0.667
	1.0E+03	4	–0.375		1.0E–03	4	0.750
	1.0E+04	5	–0.500		1.0E–02	5	0.833
DALYs per case	1.0E–04	1	0.000				
	1.0E–03	2	0.200				
	1.0E–02	3	0.400	Overall score = sum(inputs scores)			
	1.0E–01	4	0.600				
	1.0E+00	5	0.800				

The comparison between the stochastic and the ordinal scoring approach with linear and log-scaled scoring is shown in Figure 16. The results showed that, when ordinal scoring is used, the food–pathogen combinations are placed into quite broad sets of categories and their rankings have significant differences compared with the stochastic approach. The ranking using log-scaled scoring system gives more categories but shows less similarity with the reference ranking (Kendall's tau = 0.638) than the ranking obtained with the linear scoring (Kendall's tau = 0.733) where both rankings with ordinal scoring have more errors towards the top of the list (the average precision rank correlation coefficients were 0.417 and 0.462, respectively). According to the two measures of rank correlation, the ordinal scoring approach performed worse than the deterministic one. This can be attributed to the fact that the use of scores and simple sum of scores instead of a more complicated mathematical formula induced additional errors on the risk estimate. In general, the comparison showed that ordinal scoring approach has little resolution, with high risks and low risks having a high chance of being classified in the same rank.

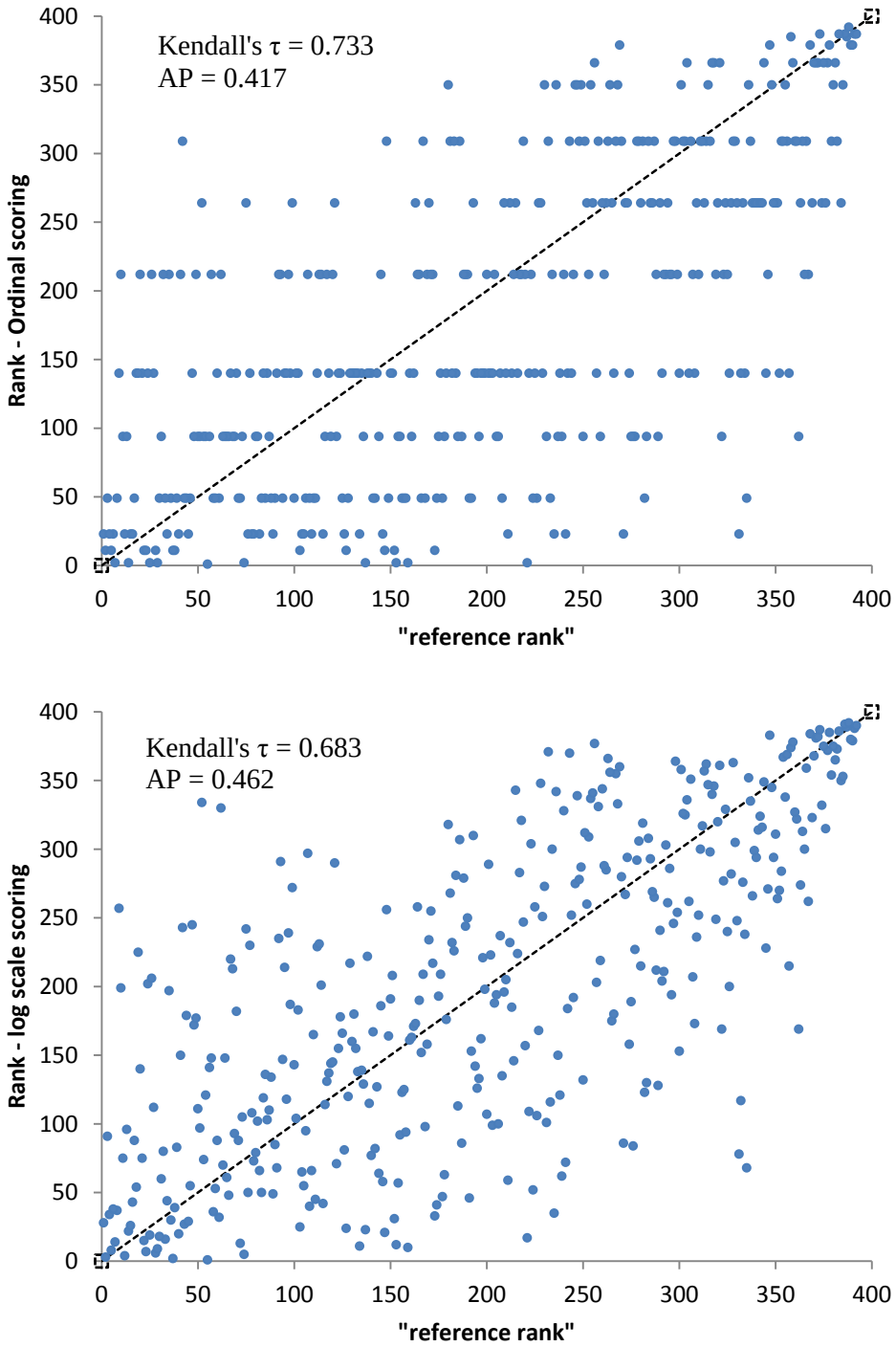


Figure 16: Comparison between the ranking obtained with the stochastic approach named as "reference rank" and rankings derived from the ordinal scoring approach with linear and log-scaled scores

4.3. Concluding remarks

The analysis performed in this section aimed at a systematic comparison of the general approaches in risk ranking (i.e. stochastic, deterministic, ordinal scoring). The results showed that both deterministic and ordinal scoring approaches may provide rankings significantly different from the stochastic approach. The difference between the deterministic and stochastic approach depends on the statistical measure used for the variable inputs. In addition, both deterministic and ordinal scoring approaches

showed more errors (i.e. differences from the stochastic approach) towards the top of the ranking list, which is important from the risk management point of view. However, of the two approaches, the deterministic one showed significantly higher similarities with the reference stochastic approach.

The use of deterministic models that ignore variability may result in risk ranking errors, which may be greater for the food–pathogen combinations with the highest risk, as shown in the example. In deterministic approaches, the selection of the point estimate used in the model can affect the risk ranking. Among different possible point estimates (arithmetic mean, median, 75th and 90th percentiles), the use of a high percentile provides, in general, ranking results which are most similar to a stochastic model. However, the performance of different point estimates in a ranking assessment will depend on the data input for the specific food–pathogen combinations involved; therefore, it is recommended to use more than one point estimates, for example arithmetic mean and a higher percentile as part of sensitivity analysis to compare rankings.

5. Uncertainty

None of the available risk ranking tools selected for this opinion is able to take into account and describe uncertainty in risk ranking. The need to characterise, document and explain uncertainty in risk assessment has been recognised by EFSA (2009). Although the number of published studies on the various methods for incorporating uncertainty in risk assessment is increasing, less information is available for risk ranking. The objective of this section is to present methodologies for identifying and evaluating the uncertainty sources in risk assessment models as well to explore their applicability to risk ranking models using a case study.

5.1. Background

In the EFSA context, the term "uncertainty" is intended to cover '*all types of limitations in knowledge, at the time it is collected*' in the risk assessment process (EFSA, 2009). The need to address uncertainty is expressed in the Codex Working Principles for Risk Analysis. These state that '*constraints, uncertainties and assumptions having an impact on the risk assessment should be explicitly considered at each step in the risk assessment and documented in a transparent manner*' (CODEX, 2007). The Scientific Committee of EFSA explicitly endorsed this principle in its guidance on transparency in risk assessment (EFSA, 2009). Therefore, it is recognised that in the risk assessment process it is important to characterise, document and explain all types of uncertainty arising in the process.

Ideally, the analysis of the uncertainty in a risk assessment would require:

- identifying uncertainties;
- describing uncertainties;
- evaluating uncertainties around individual factors in their own scales;
- evaluating the impact of individual factors uncertainties on the assessment outcome;
- evaluating the combined impact of multiple uncertainties on the assessment outcome, including evaluating how much the combined uncertainties downgrade the weight of the evidence.

The last three steps can be conducted at three levels: descriptive, deterministic and probabilistic.

An EFSA Working Group (WG)[6] is currently formulating guidelines on how the uncertainty analysis should be performed in a harmonised and structured way.

[6] See http://www.efsa.europa.eu/en/scer/scerwgs.htm for details.

5.2. Typology of uncertainty

Organisations operating at the boundary between science and policy, such as EFSA, need to address very complex issues that often involve high stakes. Dealing with uncertainties in such issues also implies a vision on the role of science in policy making (Petersen et al., 2011). Increasingly, it has become clear that science cannot be value-free and that politics need to deal with issues that are clouded with uncertainties, including value diversities. This implies that uncertainty is more than a number and can include the following dimensions:

- technical (inexactness)
- methodological (unreliability)
- epistemological (ignorance)
- societal (limited social robustness).

Communicating uncertainty to risk managers in such a way that they can adequately include different possible outcomes of the risk assessment in their decisions is a key requirement and will be further discussed in an opinion of the Scientific Committee on uncertainty, which will be published for public consultation in 2015. The present opinion will explore how uncertainty in risk ranking models can be identified and characterised, and how the impact of uncertainty on the risk ranking results can be evaluated qualitatively or quantitatively.

A typology for identifying and characterising uncertainty sources has been presented by Knol et al. (2009), see Table 44. The value of this classification is that it helps to define further actions to deal with the identified uncertainties. It also offers a framework for transparent identification and description of all uncertainties involved in a risk assessment, including aspects that have not been included in the problem formulation or system boundaries. Sources of uncertainty are related to the risk assessment question. For example, if (representative) data are available for one specific country, they would not be a source of uncertainty if the risk assessment was related to that country, but would be if the risk assessment concerns other countries.

Table 44: Typology of uncertainties (obtained from Knol et al. (2009))

Uncertainty characterisations	Categories
Location: the location at which the uncertainty manifests itself in the assessment	**Context:** definitions and boundaries of the system that is being assessed **Model structure:** structure and form of the relationships between the variables that describe the system **Parameters:** constants in functions that define the relationships between variables (such as relative risks or severity weights) **Input data:** input datasets (such as concentrations, demographic data, and incidence data)
Nature: the underlying cause of the uncertainty	**Epistemic:** resulting from incomplete knowledge **Ontic (process variability):** resulting from natural and social variability in the system
Range: expression of the uncertainty	**Statistical (range + chance):** specified probabilities and specified outcomes **Scenario (range + "what if"):** specified outcomes, but unspecified probabilities
Recognised ignorance: unknown outcomes, unknown probabilities—uncertainties are present, but no useful estimate can be given	
Methodological unreliability: methodological quality of all different elements of the assessment; a qualitative judgement of the assessment process which can based on, for example, its theoretical foundation, empirical basis, reproducibility and acceptance within the peer community	
Value diversity among analysts: potential value-ladenness of assumptions which inevitably involve—to some degree—arbitrary judgements by the analysts	

Further details on the uncertainty typology can be found in the technical report accompanying this opinion (Bouwknegt and Havelaar, 2014).

5.3. Evaluation of uncertainty sources: NUSAP approach

One approach to deal with uncertainties after its identification and characterisation is the NUSAP system (van der Sluijs et al., 2005).This provides a structured approach to appreciating uncertainties in model-based health risk assessments. NUSAP stands for numeral, unit, spread, assessment and pedigree. The first three dimensions are related to conventional technical approaches to uncertainty, expressed in numbers (N) with appropriate units (U) and a measure of spread (S) such as a range or standard deviation. Methods to address spread include statistical methods, sensitivity analysis and expert elicitation. The last two dimensions are related to aspects of uncertainty that can less readily be analysed by quantitative methods. Assessment (A) expresses qualitative expert judgements about the quality of the information used in the model. Pedigree (P) implies a multi-criterion evaluation of the process by which the information was produced. The background history by which the information was produced is considered, in combination with the underpinning and scientific status of the information. Qualitative judgements about the nature are supported by so-called pedigree matrices, which are then translated in a numerical, ordinal scale. The NUSAP output is a score per uncertainty source for the strength of the information and its influence on the model outcome. These two parameters are combined for all uncertainty sources in a diagnostic diagram, which will help to identify the key uncertainties in the assessment, i.e. those sources with a low information strength and a large influence on the model outcome. The NUSAP approach, therefore, can be used to evaluate uncertainties that cannot be quantified, but can also be useful in identifying the most important uncertainties for further quantitative evaluation and/or additional work to strengthen the evidence base of the assessment. Pedigree matrices have been developed to evaluate model parameters and input data as well as assumptions. Experts are asked to evaluate each uncertain parameter or input data and to note down the rationale for their evaluation. The strength of the information is then summarised as the median score over all experts and dimensions. However, the noted rationales are of equal importance when considering the results and the way forward.

In addition to parameters and input data, all models include a set of assumptions, which may be explicitly stated or be implicitly present in the model formulation. Identifying assumptions is a highly useful method to assess the scientific validity and credibility of model-based results. All possible assumptions should be included, e.g. processes kept out of the system boundaries, simplifications of reality, up- or downscaling in the coupling of models, embedded risk management aspects (e.g. conservative estimates), feedback loops not included, etc. A pedigree matrix for evaluating assumptions is presented in Table 45, Section 5.4. The evaluation process is similar to that for parameters. Note that Table 45 also includes a column to assess the influence on results of the assumptions; the same scale can be used for assessing the strength of model parameters.

The analysis is completed by presenting the information in diagnostic diagrams, which are presented in the next section. Parameters or assumptions with low pedigree scores (i.e. high potential value-ladeness) and high influence on results are most critical to the model and need further attention.

5.4. Case study on NUSAP to characterise uncertainty in the EFoNAO model

To evaluate the uncertainty typology and NUSAP approach, a case study was selected by evaluating the EFoNAO-RRT model, used for identifying and ranking pathogen and food combinations of most public health concern (EFSA BIOHAZ Panel, 2013). Uncertainties were identified by reviewing the approach as described in this opinion and listing explicit and implicit assumptions and uncertainties. The list of uncertainties was then finalised by discussions with experts from the WG risk ranking tools and (re)phrased as assumptions. The assumptions were subsequently characterised based on the uncertainty typology from Knol et al. (2009) (Table 46, Section 5.2). Sources of uncertainty were characterised in the following dimensions: location, nature, range, recognised ignorance, methodological unreliability and value diversity.

To identify the uncertainty sources that were most important for the total uncertainty of the EFoNAO model, the NUSAP approach was applied in a workshop involving experts from the BIOHAZ Panel and Scientific Committee. The strength of each uncertainty source was scored according to four criteria (see Table 45 for the criteria used and the scores in this study). The median of all scores for these four criteria over all experts was the measure of strength of the information. Experts also estimated the influence of the uncertainty on the model results. The median of this score, combined with the median of the strength, gives an impression of the importance of an uncertainty source: sources with low strength and large influence on the final results are most important for further consideration. The model outcomes for evaluation were (1) the identification of important microbial hazards related to foods of non-animal origin and (2) the ranking of these hazards. Note that the objective of the workshop was to evaluate the use of NUSAP in EFSA, rather than to evaluate the EFoNAO-RRT. A detailed report of the workshop is provided in Bouwknegt and Havelaar (2014).

Table 45: The pedigree matrix used in the NUSAP workshop to assess the strength of the information for each uncertain assumption and its influence on the results (effect)

	Strength				Effect
Score	**Influence of situational limitations**	**Plausibility**	**Choice space**	**Agreement among peers**	**Influence on results**
0	Choice assumption hardly influenced	The assumption is very plausible (based on established theory, verified through peer review)	Hardly any alternative available	A large majority (90–100 %) among peers of have made the same assumption	The assumption has no or negligible impact on the results
1	Limited influence in choice assumption	Plausible (based on model with theoretical basis, empirically verified data)	Very limited number of alternatives available	Many experts (75 %) would have made the same assumption	The assumption has little impact on the results
2	Choice assumption moderately influenced	The assumption is acceptable (based on a simple model, extrapolated data)	Limited choice from alternative assumptions	Several experts (50 %) would have made the same assumption	The assumption has a moderate impact on the end result
3	Important influence in choice assumption	Assumption is doubtful (based on not verified empirical data)	Average number of alternatives	Few experts (25 %) would have made the same assumption	The assumption has an important impact on the end result
4	Totally different assumption had there not been limitations	The assumption is fictive or speculative	Ample choice from alternative assumptions	Controversial assumption: hardly any experts (1 %) would have made the same assumption	

Sixteen assumptions relating to the EFoNAO-RRT were identified and analysed with the uncertainty typology (Table 46). The majority of uncertainty sources (11 out of 16) related to the parameter and input data that were used. Furthermore, 14 of the 16 uncertainties were related to imperfect knowledge ("epistemic"), which could, in theory, be reduced by further studies. These uncertainty sources resulted from the study boundaries set by the mandate, or from analysts' or technical constraints (data availability, limits in modelling techniques, etc.).

Table 46: Characterisation of the 16 uncertainty sources by using the uncertainty typology of Table 44

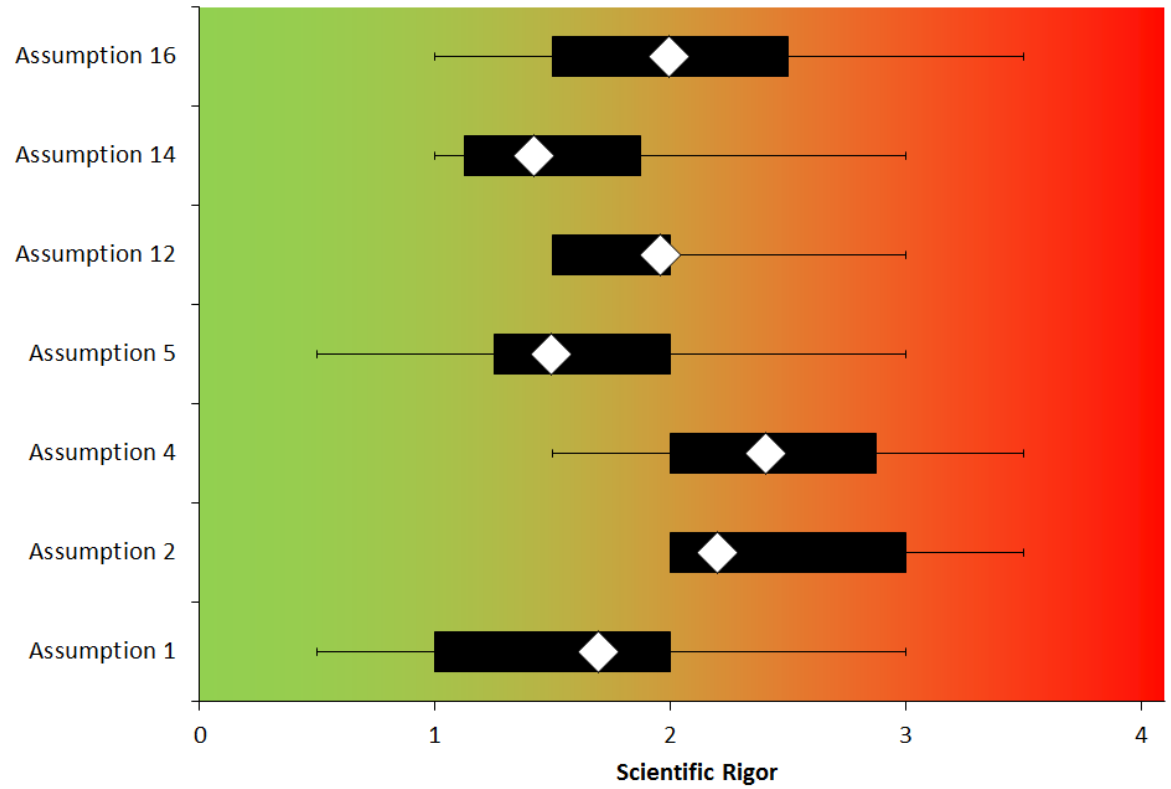
	Nature		Range		Recognised ignorance	Method unreliability	Value diversity
	Epistemic	Ontic	Statistical	Scenario			
Contextual uncertainty							
Link between a pathogen and a type of EFoNAO can be deduced from outbreak data only	X			X	+	+	+
The added value of considering pathogen inactivation to assess risk levels is negligible for each food–pathogen pair	X	X		X	+	–	+
Contextual and model uncertainty							
The risk of a pathogen/food combination can be estimated by a linear, unweighted combination of scores on seven parameters, each divided in three or four categories that are represented by arbitrary numbers	X			X	+	–	+
Model uncertainty							
The risk of a pathogen/food combination can be estimated by a combination of top-down and bottom-up approaches	X				+	–	+
Assuming a prevalence score of 2 to the category defined as "unknown prevalence", implies that the prevalence cannot be assumed to be zero for *Shigella* spp., *Yersinia* spp., *Staphylococcus aureus*, Norovirus, hepatitis A virus (HAV) and *Cryptosporidium* spp.	X			X	+	–	+
Parameter and input data uncertainty							
The estimated true number of illnesses by a specific pathogen in the EU, without consideration of attribution to sources, is a valid indicator of the risk of a specific pathogen in a specific food of non-animal origin	X			X	+	–	+
The prevalence of pathogens in all EFoNAO samples is a valid estimate for the prevalence in the EFoNAO group under consideration	X		X		+/–	–	+
The relative degree of underreporting of outbreak cases is the same in the USA and EU and for each food–pathogen pair	X		X		+	+	+
The incidence of Norovirus and bacterial intoxications in the EU is similar to the Netherlands	X		X		+	–	+
The longest reported shelf life of food in a specific food group is representative of all products in that group and pathogen growth is not affected by growth of spoilage organisms		X	X		–	–	+

Table 46: Characterisation of the 16 uncertainty sources by using the uncertainty typology of Table 44 (continued)

	Nature		Range		Recognised ignorance	Method unreliability	Value diversity
	Epistemic	**Ontic**	**Statistical**	**Scenario**			
Available consumption data are representative of the whole EU	X		X		+	–	+
Low numbers of *Salmonella* spp., *Shigella* spp., STEC and *Yersinia enterocolitica* can cause disease without growth during storage in retail or consumer's homes	X			X	+	–	+
Pathogen-specific DALY estimates published for the Netherlands are representative for the whole EU	X		X		–	+	+
DALYs per case for *Shigella* spp. and *Y. enterocolitica* fall within the same category as *Salmonella* spp. and are the same for STEC O157 and STEC non-O157	X		X		+	–	+
All products will be eaten at the end of their shelf life		X	X		–	–	+
With the exceptions of *Bacillus cereus* and *Clostridium perfringens*, the overall prevalence of all pathogens in the different EFoNAO groups, is assumed to be either low (< 1 %) or unknown	X		X		–	+/–	+

DALY: disability-adjusted life years; EFoNAO: EFSA food of non-animal origin; STEC: Shiga toxin-producing *Escherichia coli*.

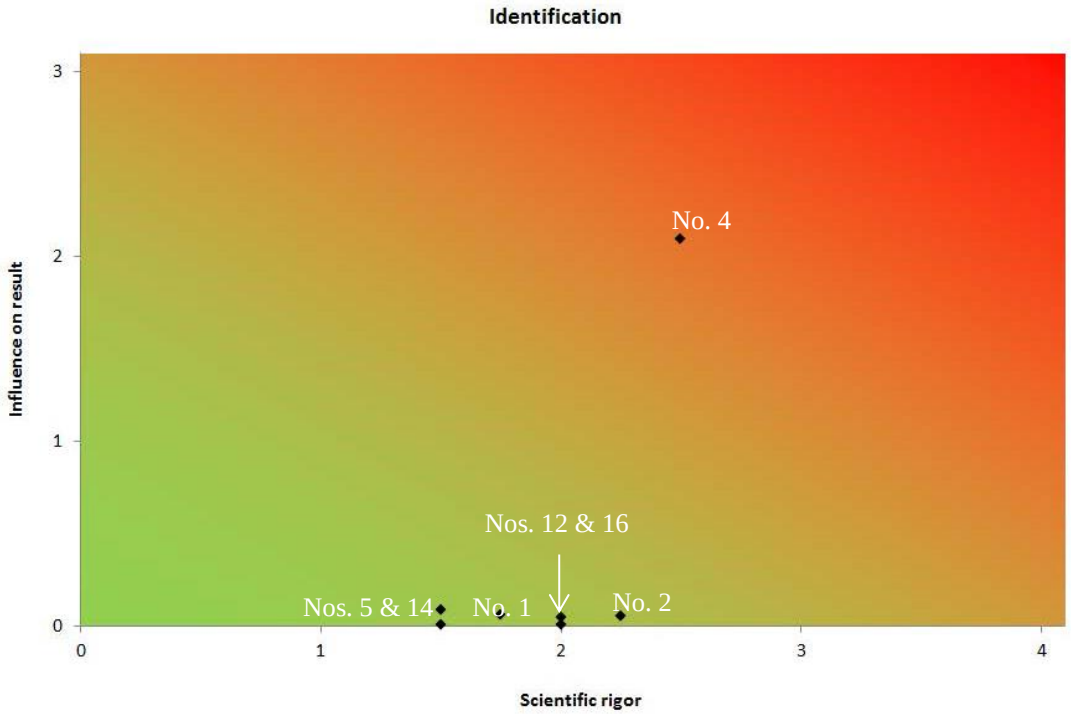
During the workshop, there was intensive discussion on the NUSAP methodology and on the interpretation of the criteria and the scores. This discussion led to a revision of the definitions for the "influence on results" categories, as reflected in Table 46. As a consequence of the time needed for these discussions, only 7 out of 16 assumptions could be evaluated. Figure 17 shows the scoring results for the strength of the information for these seven assumptions. Six out of seven sources had an interquartile range covering two score classes, thereby showing agreement among most experts. However, the range of scores covered the full scale (four classes) for three assumptions and three classes for four assumptions, indicating that for all assumptions, opinions diverging from the majority view were expressed. One uncertainty source (no 16, scored first of all assumptions), had an interquartile range covering three classes. The median scores are concentrated around the midpoint of the scale, which may reflect the divergence in scores by individual expert and may be related to lack of experience of the experts.



The white diamonds indicate the median score, the error bars indicate the minimum (left) and maximum (right) score and the black rectangles indicate the interquartile range. Assumptions with higher scores (in the red zone) have lower strength of the information compared with lower scores (green zone). Diamonds crossing the y-axis indicate the assumptions that have not been scored.

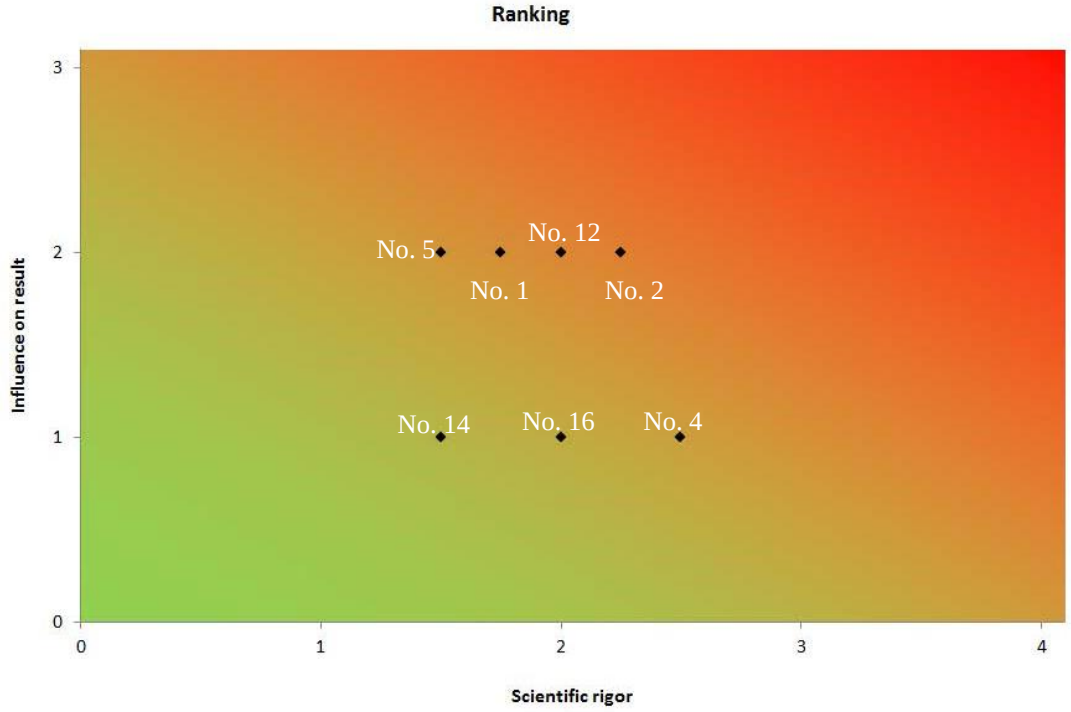
Figure 17: Strength (scientific rigor) of the information for the assumptions identified in the EFoNAO-RRT that yield uncertainty in the model outcome

Figures 18 and 19 show the strength and effect diagrams for the model outcomes hazard identification and hazard ranking. Only assumption "No 4" was judged to be influential on the hazard identification, whereas four assumptions (Nos 1, 2, 5 and 12) were judged to have a moderate impact on hazard ranking.



The x-axis displays the median strength of the information (i.e. the white diamonds from Figure 18), the y-axis the median score for the influence on results. Values with a high score on influence on results and strength of the information (the red zone) are critical assumptions in the model.

Figure 18: Strength (scientific rigor) and effect diagram for the seven assumptions of the EFoNAO-RRT scored during the workshop for the influence on hazard identification



The x-axis displays the median strength of the information (i.e. the white diamonds from Figure 19), the y-axis the median score for the influence on results. Values with a high score on influence on results and strength of the information (the red zone) are critical assumptions in the model.

Figure 19: Strength (scientific rigor) and effect diagram for the seven assumptions of the EFoNAO-RRT scored during the workshop for the influence on hazard ranking

The combination of uncertainty typology and NUSAP was found to be very helpful by all participants of the workshop. The procedure helped to systematically identify and evaluate the uncertainty sources related to model outcomes and to assess their impact on the end results. A framework encompassing uncertainty typology and evaluation (e.g. by NUSAP) should be part of each risk assessment to formalise discussions on uncertainties. By doing this structurally and integrated with the risk assessment activities, experience grows and the process would eventually save time. Nevertheless, it was recommended that practicality and feasibility aspects should always be considered when incorporating uncertainty assessment in the risk assessment process.

The interpretation of pedigree criteria to assess the strength and effect (see Table 47) was found to be difficult by participants. Part of the difficulty is caused by the difference in terminology used by scientists working in philosophical sciences, who developed the NUSAP methodology, and those working in the natural sciences. It was recommended that a clear terminology is developed, which is understood by all involved in the assessment. Preferably, a short training session with dummy uncertainty sources would be conducted before the NUSAP workshop.

Much time during the workshop was devoted to discussions on how to describe the sources of uncertainty. Ideally, consensus on the phrasing/wording is obtained before the scoring starts. This should be an iterative process that involves both the principal analyst(s) of the study to be evaluated and the experts who will participate in the NUSAP workshop.

The aggregation of scores by all experts on all four criteria related to the strength of the information of the assumptions in a single median (and a range around it) was considered to result in loss of information. In the final report (Bouwknegt and Havelaar, 2014), the scores were also presented by criterion. The pedigree criterion “agreement among peers” was scored consistently best for all assumptions; the criteria “influence of situational limitations”, “plausibility” and “choice space” scored, in general, lower than “agreement among peers” and showed larger variation. The criteria considered in the scoring of strength of the information are different in nature and addressing potential issues may require different strategies. A more detailed summary description of the results of this analysis, and possible the development of a multi-criterion analysis within the NUSAP approach was proposed in order to extract and use more of the information obtained during scoring of all criteria.

5.5. Quantifying uncertainty in risk ranking

In practice, the parameters and data used in risk ranking cannot be characterised precisely; the knowledge of the causal phenomena and available data are generally incomplete. Such uncertainty propagates within the model and causes variability in its outputs; as many values are possible for a model parameter, the model outputs associated to the different values of the uncertain parameter will be different. Following a qualitative analysis (e.g. by the NUSAP method), the quantification and characterisation of the resulting output uncertainty is crucial, and it defines the scope of the uncertainty analysis. Such quantitative analysis could be initially focused on or even be restricted to those parameters that are considered most influential on the model outcomes by the qualitative analysis.

5.5.1. Principles of uncertainty analysis

Uncertainty analysis consists of evaluating quantitatively the uncertainty or variability in the model components (parameters, input variables, equations) for a specific situation, and generating an uncertainty distribution for each output variable instead of a misleading single value. An important consequence is that it provides tools to assess, for instance, the probability of one food–pathogen combination is at higher risk than another combination. This makes uncertainty analysis a key component of risk ranking.

Within a particular model, equations, parameters and input variables are all subject to variability or uncertainty. First, decisions have to be made on the model structure and on the functional relationships between input variables and output variables. These decisions may sometimes be somewhat

subjective, and it is not always obvious what their consequences will be. Thus sensitivity analysis needs to be performed to establish the effects of one or several type of modelling approaches on the output of the model. Second, parameter values are obtained from statistical estimation procedures based on empirical evidence or sometimes from literature reviews or expert opinion. Their quality is inevitably limited by the variability and possible lack of appropriateness of the available evidence. The uncertainty and natural variability of parameters are the central point of many uncertainty analyses.

For each input, the uncertainty needs to be defined. The uncertainty can be described in different ways. For a parameter, it is often given as the most likely value plus or minus a given percentage or it is specified through a continuous probability distribution over a range of possible values. In general, three characteristics may be considered for describing the uncertainty: nominal values, uncertainty domains and probability distributions. The uniform distribution, which gives equal probability to each value within the uncertainty range, is frequently used in sensitivity analysis when the main objective is to understand model behaviour. In uncertainty analysis, more flexible probability distributions are usually needed to represent the input uncertainty.

In practice, uncertainty analysis consists of four steps:

- definition of the distribution of each uncertain input factor;
- generation of N iterations from the distribution of uncertain input factors;
- computation of the model output for each set of iterations;
- analysis of the output distributions (computation of means, variances, quartiles, percentiles, etc.).

The first step of an uncertainty analysis is to define the probability distributions for the input parameters. Table 47 gives an example of a risk assessment model inputs. A risk assessment model usually describes the variability of the occurrence of a list of events using stochastic processes. It is crucial to distinguish variability probability distributions from uncertainty distribution parameters. Attention must be paid when choosing probability distributions. The range of input values usually has more influence on the output than the distribution shapes, but some characteristics such as the degree of symmetry or skewness may also play a role.

There is a large choice of probability distributions. The uniform distribution gives equal weight to each value in the uncertainty range. However, the extreme values of the uncertainty ranges are less likely than the central values and other distribution are needed. The well-known normal distribution, a symmetrical distribution, is often convenient since it requires only the specification of two well-understood parameters: a mean value and a standard deviation. For some inputs, the distribution should be asymmetrical, for example if the input is greater than zero. Then log-normal, gamma or beta distributions offer a large range of possibilities. In uncertainty analysis, normal distribution is often replaced by the truncated normal distribution or by symmetric beta distributions, which give upper and lower bounds to the possible values. Finally, the triangular distributions are often convenient for a simple representation of subjective beliefs, because they are defined entirely by their uncertainty range (minimum and maximum) and their most likely value.

The uncertainty analysis brings additional challenge to risk ranking. In Figure 20, the uncertainty probability density distributions of the model outputs for risks associated with three food–pathogen combinations are presented. The rankings depend on what statistic is used to characterise a risk whose value is not known with certainty. If means are used (as a best guess) the three food–pathogen combinations would be ranked 2-1-3. If, for example, 99th percentiles are used (as a worst case), the order becomes 3-2-1.

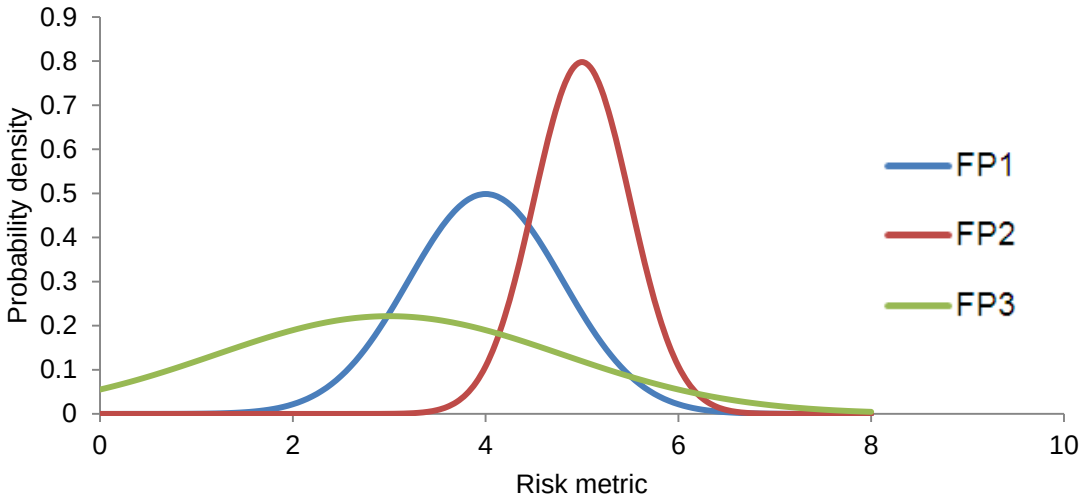


Figure 20: Uncertainty probability density distributions of the model outputs for risks associated with three food–pathogen (FP) combinations. Note that the x-axis uses a hypothetical risk metric, therefore no units are used.

5.5.2. Ranking in presence of uncertainty

First, a method is presented for comparing the risk associated with two food–pathogen combinations. For example, we have two combinations A and B in presence of uncertainty on the parameters used to assess the associated risk for consumers, which propagate through the model leading to uncertainties in risk estimates. In this case, risk calculations should reflect these uncertainties and so should the ranking. For simplicity of illustration, log-normally distributed uncertainty is assumed to be affecting directly the risks for A and B.

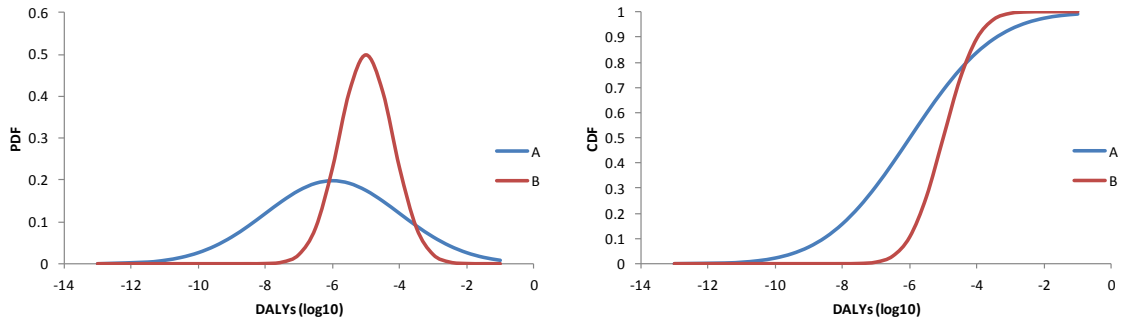


Figure 21: Probability density functions (PDF) and cumulative distribution functions (CDF) of the random variables DALYs for A and B. Above the 75th percentile B > A, below the 75th percentile A > B.

Examining the distributions of the DALYs associated with A and B in Figure 21, distribution A (DA) and distribution B (DB), respectively, one may observe that the DA is much more uncertain than DB but the expected value of DB is greater than DA. On the other hand, there is a range in which the DB percentiles are larger than the DA ones. For example, if one were to perform the ranking based on the DALYs 95th percentile values, the conclusion would be that combination A is more risky than B, contrary to what would happen if the rankings were based on the expected values.

The drawback of comparing the expected values or specific percentiles lies in the loss of information about the distribution. In order to give full account of the difference between the distributions of DA and DB one have to consider the random variable DA–DB whose PDF and CDF are shown in Figure 22.

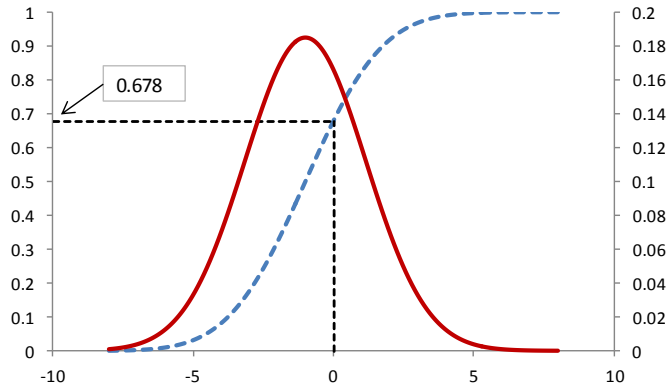


Figure 22: PDF and CDF of the random variable (DA–DB). The probability of DA–DB < 0 is 0.678.

In order to establish whether A is more risky than B, one can consider the probability rAB = 1 – P(DA–DB < 0) (0) that DA is greater than DB; for example, in the present case rAB = 1 – 0.678 = 0.322, which means that, with a probability of 0.322, A is more risky than B. To decide on the relative importance of the two combinations A and B, one may choose a threshold (T) ranging from 0.5 to 1 on the rAB value such that, if rAB is larger than T, then A is more risky than B, otherwise no conclusion can be drawn. Obviously, the lower the threshold, the higher the risk associated with the decision. However, the choice of a simple-valued threshold has some limitations when considering multiple combinations. These limitations can partially be overcome by referring the comparison to a threshold range [Tl, Tu] in such a way that for the two components A and B (Baraldi et al., 2009):

- if rAB > Tu, then A is more risky than B;
- if rAB < Tl, then B is more risky than A;
- if Tl < rAB < Tu, then A is equally risky to B.

To extend the method to systems with a large numbers of components, a procedure for successive ranking must be introduced to avoid the combinatorial explosion of pairwise comparisons using, for example, the Quicksort algorithm (Horae, 1962) implemented by Baraldi et al. (2009). Once the probability distributions have been specified, representative samples are drawn from these distributions using Monte Carlo sampling. The samples are drawn independently, and each sample is generated by drawing independently the value of each parameter.

After the sample of parameters values have been generated, the corresponding model output values are computed. If the computation of the model output is time consuming, this step may be difficult to carry out. In this case, the sample size (N) must be changed to a smaller value because of the computation time.

The last step of the analysis is to summarise the values of obtained outputs. Different quantities can be easily calculated. For example, when the model has a single output variable, estimates of the expected value and variance of can be computed. It is also useful to estimate the quartiles/percentiles associated with the distribution and the probabilities that the output variable is lower than some thresholds. A histogram representation of the output variable values can also provide more information than the summary statistics.

5.5.3. Example of uncertainty analysis

For this example, the generic assessment framework presented in Section 4 is applied to five food–pathogen combinations. First, the stochastic quantitative risk assessment, including only variability, was applied using inputs presented in Table 47.

Table 47: Inputs for hypothetic five food–pathogen (FP) combinations

Parameters	FP1	FP2	FP3	FP4	FP5
P	0.03	0.20	0.01	0.01	0.30
m_0	1	–2	–3	–2	1.5
s_0	0.8	0.5	0.5	0.5	0.8
Growth	1	1	1	1	0
m_g	0.5	0.3	0.5	0.25	–
s_g	0.8	0.8	0.5	0.1	–
m_c	–3	–6	– 6	– 6	–5
s_c	0.8	0.5	0.5	0.5	0.5
m_s	30	100	125	125	150
s_s	7	10	5	5	10
RTE	1	1	0	0	0
m_r	–	–	– 3	–3	–7
s_r	–	–	0.5	0.5	0.5
R	–10	–10	–3	–3.5	–3
DALY	14	14	10	1	0.03
FR	30	30	50	50	80
LogR ignoring uncertainty	3.42	3.81	2.14	0.16	2.56
Rank orders	2	1	4	5	3

RTE: ready-to-eat; DALY: disability-adjusted life years; FR: frequency of consumption.

As shown in Table 48, FP2 is ranked first, FP1 is ranked second, FP5 is ranked third, FP3 is ranked fourth and FP4 is ranked fifth. This ranking assumed that all used parameters are certain (known perfectly).

In practice, parameters cannot be estimated precisely; as the knowledge of the causal phenomena and available data are generally incomplete. In Table 48, the choice of the types of probability distributions is presented. The ranges of the possible values of the parameters for the five hypothetical food–pathogen pairs are shown in Figure 23. The choice of the types of probability distributions and their parameters can be based on available data using classical inferential statistical approaches or statistical Bayesian approaches, or obtained from formal expert elicitation knowledge exercises.

Table 48: Variability and uncertainty probability distributions

Model inputs	Parameters of the variability distribution	Variability distribution model (first order iteration in Figure 24)	Uncertainty distribution (second order iteration in Figure 24)
Prevalence	p	Bernoulli (p)	p ~ beta (a, b)
Initial concentration in $\log_{10}$ CFU/g	m_0 s_0	Normal (m_0,s_0)	M_0 ~ normal (x, y) S_0 ~ gamma (z, w)
Growth potential in $\log_{10}$	m_g s_g	Gamma (m_g, s_g)	m_g ~ normal (t, u) s_g ~ gamma (d, f)
Cross-contamination ($\log_{10}$ probability of transfer)	m_c s_c	Normal (m_c,s_c)	m_g ~ normal (q, s) s_g ~ gamma (g, h)
Portion size	m_s s_s	Gamma (m_s, s_s)	m_g ~ normal (k, l) s_g ~ gamma (n, r)
Potential reduction	m_r s_r	Normal (m_r,s_r)	m_g ~ normal (i, o) s_g ~ gamma (p, m)
Dose–response	r	No variability	p ~ beta (a', b')
DALY	DALY	No variability	DALY ~ gamma (v,e)
Frequency of consumption	FR	No variability	FR ~ normal (j,k')

DALY: disability-adjusted life years.

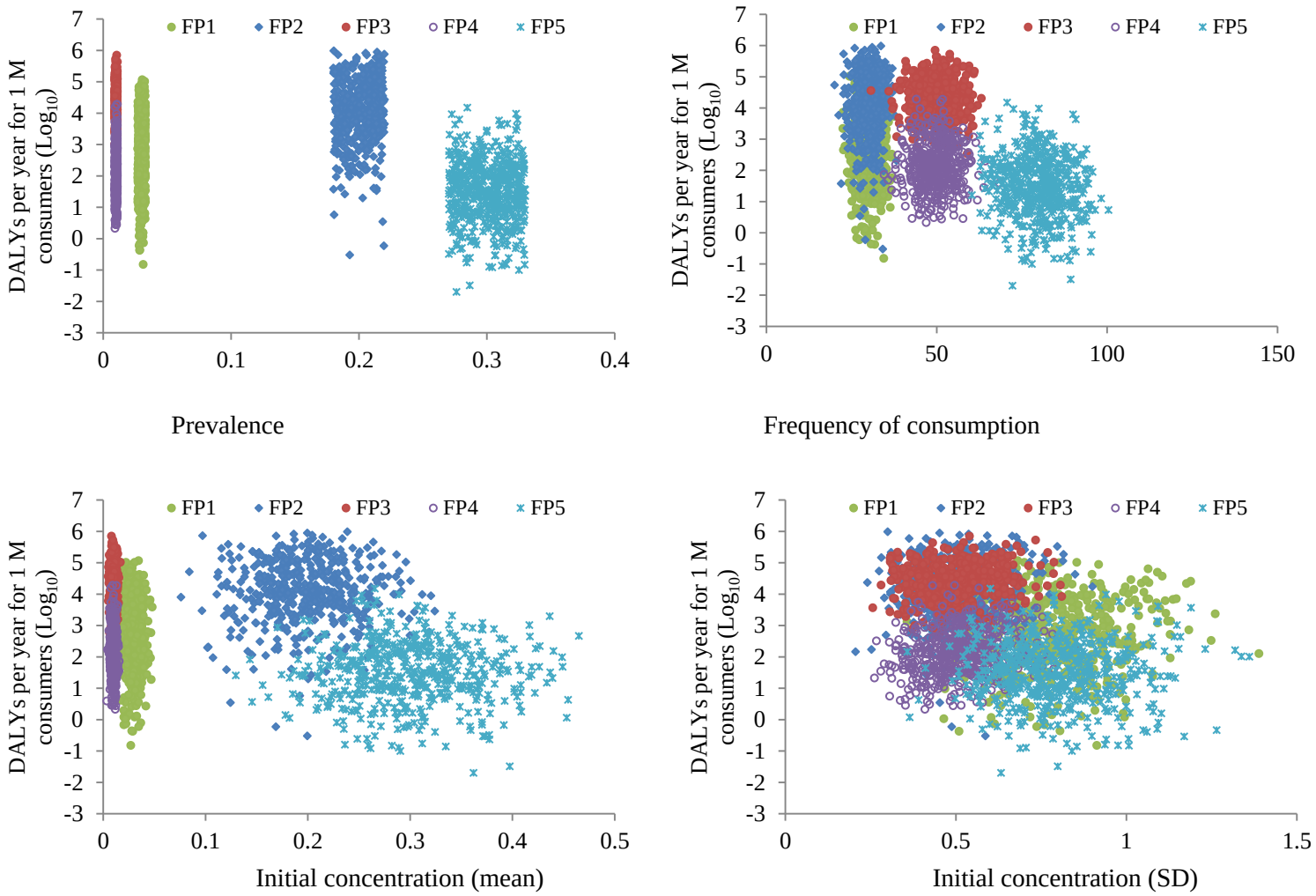


Figure 23: Uncertainty ranges of input parameters and their corresponding model outputs for the five hypothetical food–pathogen combinations (FP1 to FP5)

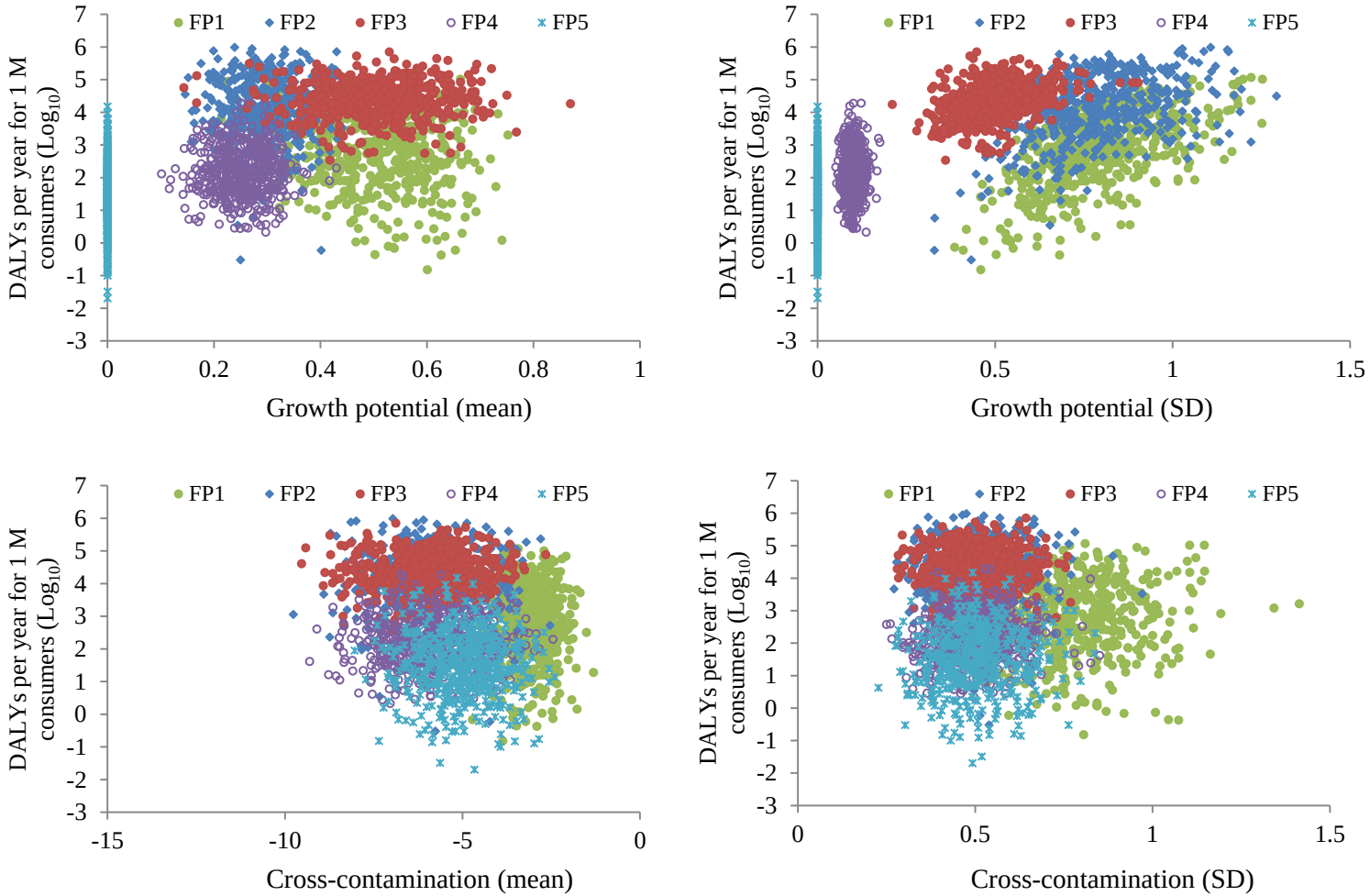


Figure 23: Uncertainty ranges of input parameters and their corresponding model outputs for the five hypothetical food–pathogen combinations (FP1 to FP5) (continued)

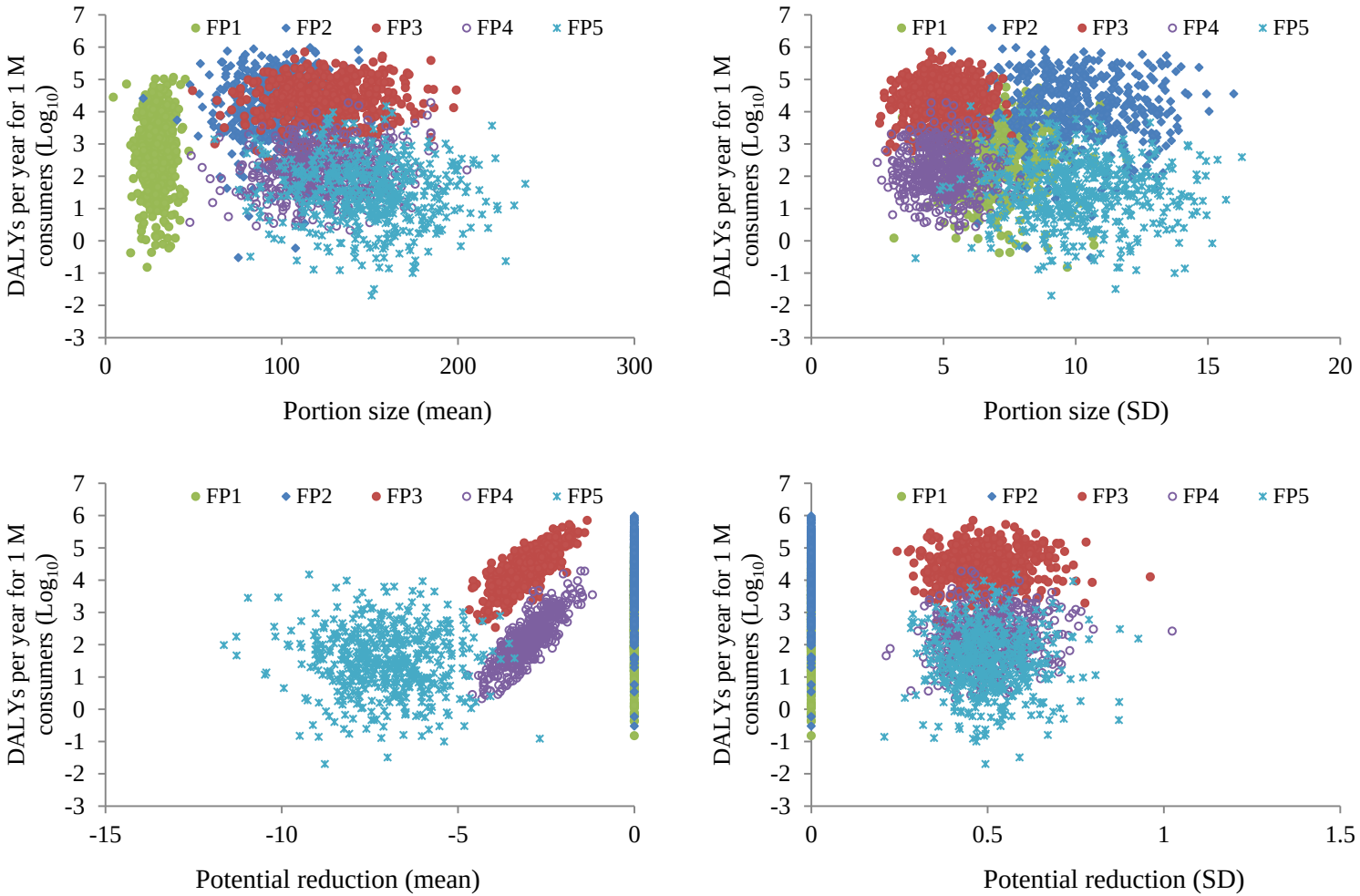


Figure 23: Uncertainty ranges of input parameters and their corresponding model outputs for the five hypothetical food–pathogen combinations (FP1 to FP5) (continued)

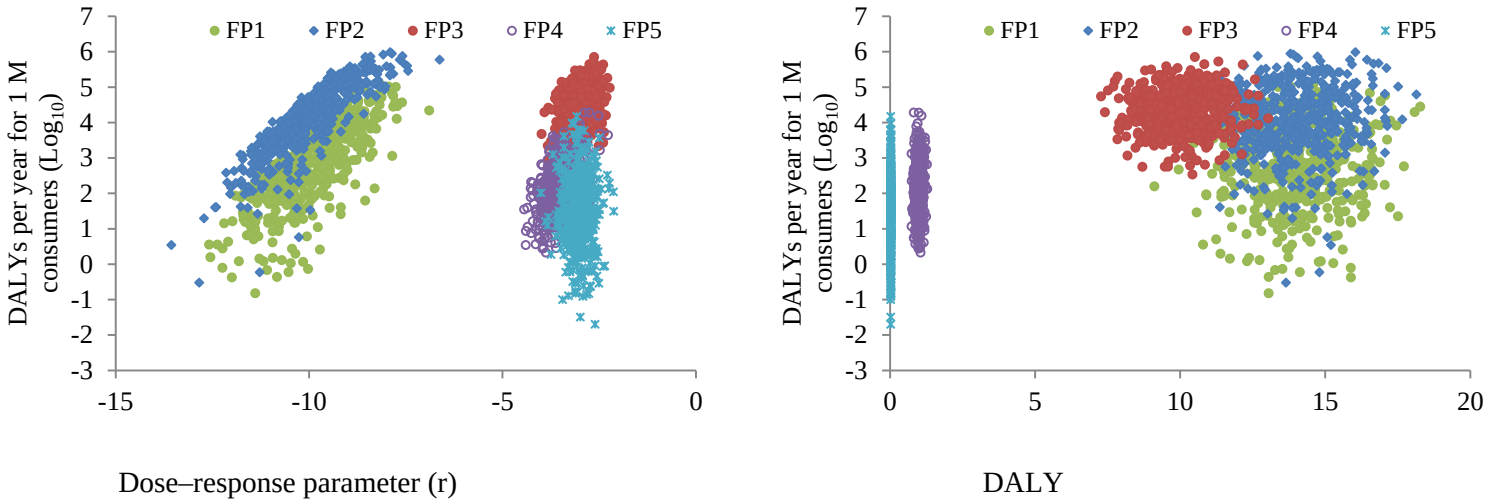


Figure 23: Uncertainty ranges of input parameters and their corresponding model outputs for the five hypothetical food–pathogen combinations (FP1 to FP5) (continued)

The uncertainty about the parameters propagates within the model and causes variability in its outputs: as many values are possible for a model parameter, the model outputs associated with the different values of the uncertain parameter will be different. Figure 24 shows how uncertainty can be integrated in the generic risk assessment framework presented in Section 4.1.1 (Figure 10 and Table 39). Two hierarchical loops are defined, an outer loop for uncertainty (second order iteration in Figure 24) and an inner loop for variability (first order iteration in Figure 24): two-dimensional (2D) Monte Carlo simulations.

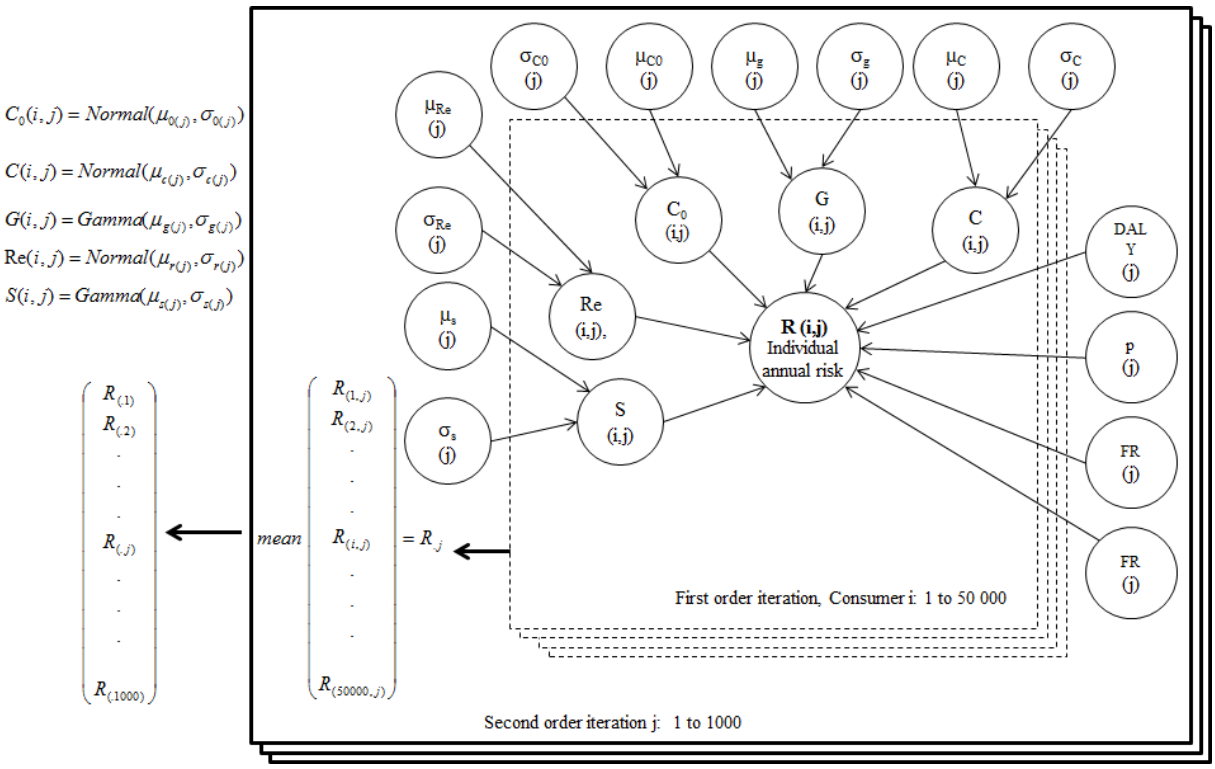


Figure 24: Structure of a risk assessment model using two-dimensional Monte Carlo simulations

The model outputs are presented in Figure 23.

In Figure 25, the median and 95th percentiles of the risks obtained for each food–pathogen pairs are presented (bar chart with different colours). When uncertainty is included, FP3 is ranked first, FP2 is ranked second, FP1 is ranked third, FP4 is ranked fourth and FP5 is ranked fifth. The other bar diagrams (with blue bars) in Figure 25 show how the order of the different food–pathogen pairs varies because of the uncertainty about the parameters used in the two-dimensional Monte Carlo simulations. From the distribution of the different ranks we assessed the probability of each food–pathogen pair being ranked higher than the other food–pathogen pairs (pairwise comparison).

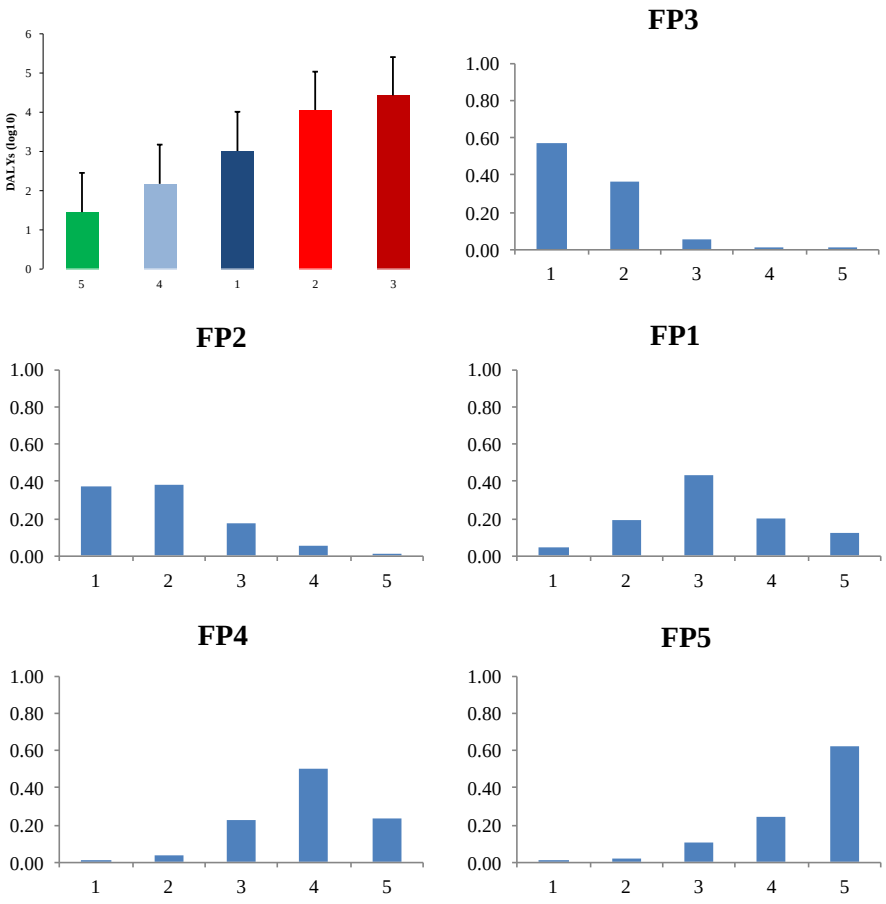


Figure 25: Probability distribution of the food–pathogen combination rank orders

Table 49: Pairwise comparison using the probability of one food pathogen being ranked higher than the other food–pathogen pairs (rAB)

Rank orders without uncertainty	Median rank (2D Monte Carlo)	FP3	FP2	FP1	FP4	FP5	Final rank (2D Monte Carlo)
4	FP3	–	0.61	0.91	0.99	0.99	FP3, FP2 (1)
1	FP2		–	0.78	0.93	0.96	FP3, FP2 (1)
2	FP1			–	0.72	0.81	FP1 (3)
5	FP4				–	0.70	FP4, FP5 (4)
3	FP5					–	FP4, FP5 (4)

Using the rule 'if rAB > 0.80, then A is more risky than B; if rAB < 0.20, then B is more risky than A; if 0.20< rAB < 0.80, then A is equally risky to B', then FP3 and FP2 have the same rank of first, FP1 is ranked third, FP4 and FP5 have the same rank of fourth (Table 49).

In summary, uncertainty in rank orders cannot be formally quantified using qualitative or semi-quantitative ranking methods even though these are often applied in situations where data are limited. Uncertainty and variability in risk ranking can be represented by means of probability distributions, for example using two-dimensional Monte Carlo simulations. However, probabilistic representation is difficult when sufficient data are not available for statistical analysis. Expert elicitation procedures to incorporate diffuse information into the corresponding probability distributions may be adopted.

6. Risk ranking toolbox for the EFSA BIOHAZ Panel

The risk ranking questions for the BIOHAZ Panel vary widely in nature and there are different constraints in time, resources and available data that need to be taken into account when deciding on a risk ranking approach. Therefore, no single solution will satisfy all needs of the Panel, and this section proposes several approaches that can be considered by the Panel for a particular mandate. These include a bottom-up approach (i.e. based on exposure data, and dose–response relationships), a top-down approach (i.e. based on disease incidence and attribution data) and a combined approach.

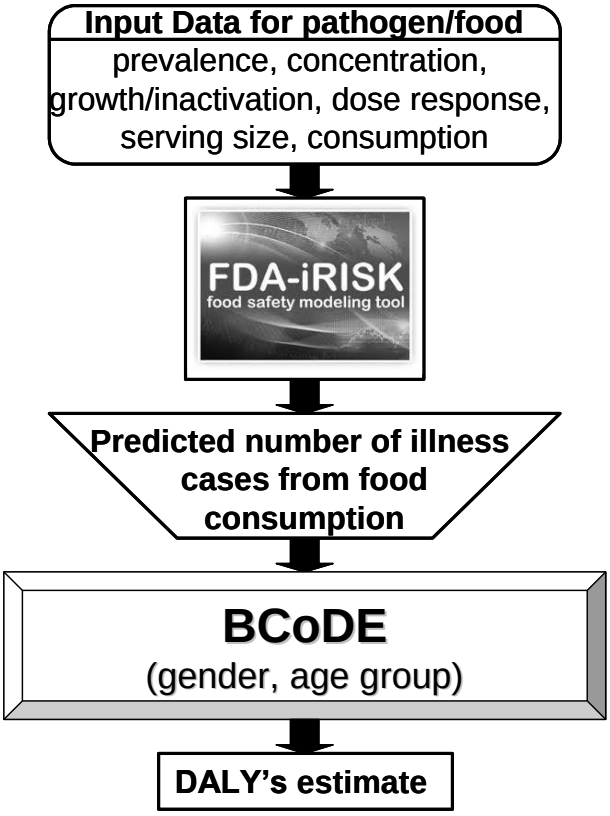
6.1. Bottom-up approach

6.1.1. Fully quantitative risk ranking

Among the tools that use a bottom-up approach for risk ranking evaluated in this opinion, the FDA-iRISK model has been identified as the most appropriate for the needs of EFSA BIOHAZ Panel. FDA-iRISK is a technically valid, fully quantitative tool providing meaningful risk metrics. It takes into account the main factors affecting risk and follows the risk assessment paradigm, while respecting the laws of probability and calculus. The FDA-iRISK tool has the best user interface among the tested tools in this opinion. Data in the correct format need to be estimated outside the tool, e.g. amount of growth, but can be entered and used easily in the tool. It is possible to model different steps in the food chain from farm to fork, providing flexibility in choosing different scenarios combining hazards, consumption patterns and processing stages. In addition, each model run can be saved and shared online with other users allowing effective quality assurance evaluation and comparison of different scenarios.

The results of the evaluation of the FDA-iRISK (version 1.0, released October 2012) conducted in Section 3 were discussed with the FDA, with a special focus on the problems related to ignoring the maximum population density and the lack of separation between variability and uncertainty, which may significantly affect the risk ranking. The tool has been further developed to add a number of new features. These features include: setting growth of the pathogens with the maximum population density, an improved treatment of rare events, the ability to perform sensitivity analysis and the addition of new distribution options (e.g. beta, beta general, truncated normal). The development of a two-dimensional Monte Carlo simulation option for the explicit separation of uncertainty and variability, and a multi-food exposure assessment feature to characterise chronic exposure is planned for a subsequent update of the tool. Version 2.0 of FDA-iRISK is expected to be released early in 2015 and the above new features are expected to significantly improve the utility of the tool.

The present opinion proposes a further improvement of FDA-iRISK outputs by combining the FDA-iRISK tool with the BCoDE tool for more detailed calculation of DALYs. In this approach, the output of the bottom-up FDA-iRISK tool expressed in total number of illnesses per year for a pathogen/food combination is used as an input in the BCoDE tool. The combination of the two tools leads to more realistic and detailed DALY estimates since it takes into account the variability and uncertainty of all variables, including gender and age group. A schematic representation of the combination of FDA-iRISK with BCoDE is shown in Figure 26.



BCoDE: ECDC Burden of Communicable Diseases in Europe (toolkit); DALY: disability-adjusted life years.

Figure 26: Combination of FDA-iRISK output and BCoDE tool for a more effective estimation of DALYs

microHibro can be considered as a future alternative option for risk ranking application by EFSA. As mentioned before, microHibro was initially developed as a microbial growth prediction tool, but with recent developments the model can be used for quantitative risk assessment and risk ranking. It has an advanced user interface and the user can design any step in the food chain from farm to fork allowing for effective data management and analysis of different scenarios combining hazards, consumption patterns and processing stages. However, since the function for risk ranking applications is in progress, the tool should be re-evaluated after completion of these developments.

sQMRA was also evaluated as a technically valid, fully quantitative, bottom-up risk assessment tool. The main weakness of sQMRA is the Excel spreadsheet format, which makes file management very complex and quality assurance and comparison of different scenarios difficult. In case of the development of a more advanced interface of the tool in the future, it could be also considered as an alternative option for the BIOHAZ Panel.

6.1.2. Semi-quantitative risk ranking: deterministic and ordinal scoring approach

The examples tested in this opinion showed that the output of deterministic models for risk ranking that do not take into account the variability of the input parameters can be significantly different compared with the reference stochastic model. In addition, the results showed that the selected statistical value of input data (arithmetic mean, median, 75^{th} and 90^{th} percentiles) does affect the risk ranking. It is concluded that, as a general principle, among the different statistical values, the use of a high percentile provides the closest ranking to the stochastic reference model. The performance of different point estimates in a ranking assessment will depend on the specific food–pathogen combinations involved. However, in all cases, the ignorance of variability in the deterministic approaches resulted in significant ranking errors, which are higher towards the top of the list when

compared with the errors towards the bottom of the list. The latter is of great importance since the top of the ranking list is more significant from the risk management point of view. Therefore, it is recommended that more than one point estimate (e.g. arithmetic mean and a higher percentile) should be used as part of sensitivity analysis to compare rankings when using deterministic models.

The opinion confirms the significant limitations of risk ranking models that use a semi-quantitative approach with ordinal scoring. Indeed, the performance of ordinal scoring models was found to deviate more from the stochastic reference model than did the deterministic models. In particular, when ordinal scoring is used, the food–pathogen pairs are placed into quite broad sets of categories and have huge risk differences compared with the reference quantitative stochastic model. The ranking using a log-scaled scoring system gives more categories but shows less similarity with the actual ranking than the ranking obtained with the ordinal scoring. Both rankings with linear and log-scaled ordinal scoring systems have more errors towards the top than the bottom of the list.

In conclusion, semi-quantitative tools based on deterministic and ordinal scoring approach may lead to erroneous risk ranking. When no other options are feasible because of limitation of available data the deterministic approach, using a high percentile of actual data for the input parameters should be preferred since it showed significantly better performance than the ordinal scoring approaches. In the case of lack of data required for a stochastic or deterministic approach, an expert elicitation should be considered. In addition, decision trees should be used only as a tool for showing how decisions about classifying food–pathogen combinations into broad categories are made (e.g. inclusion/exclusion; high/low).

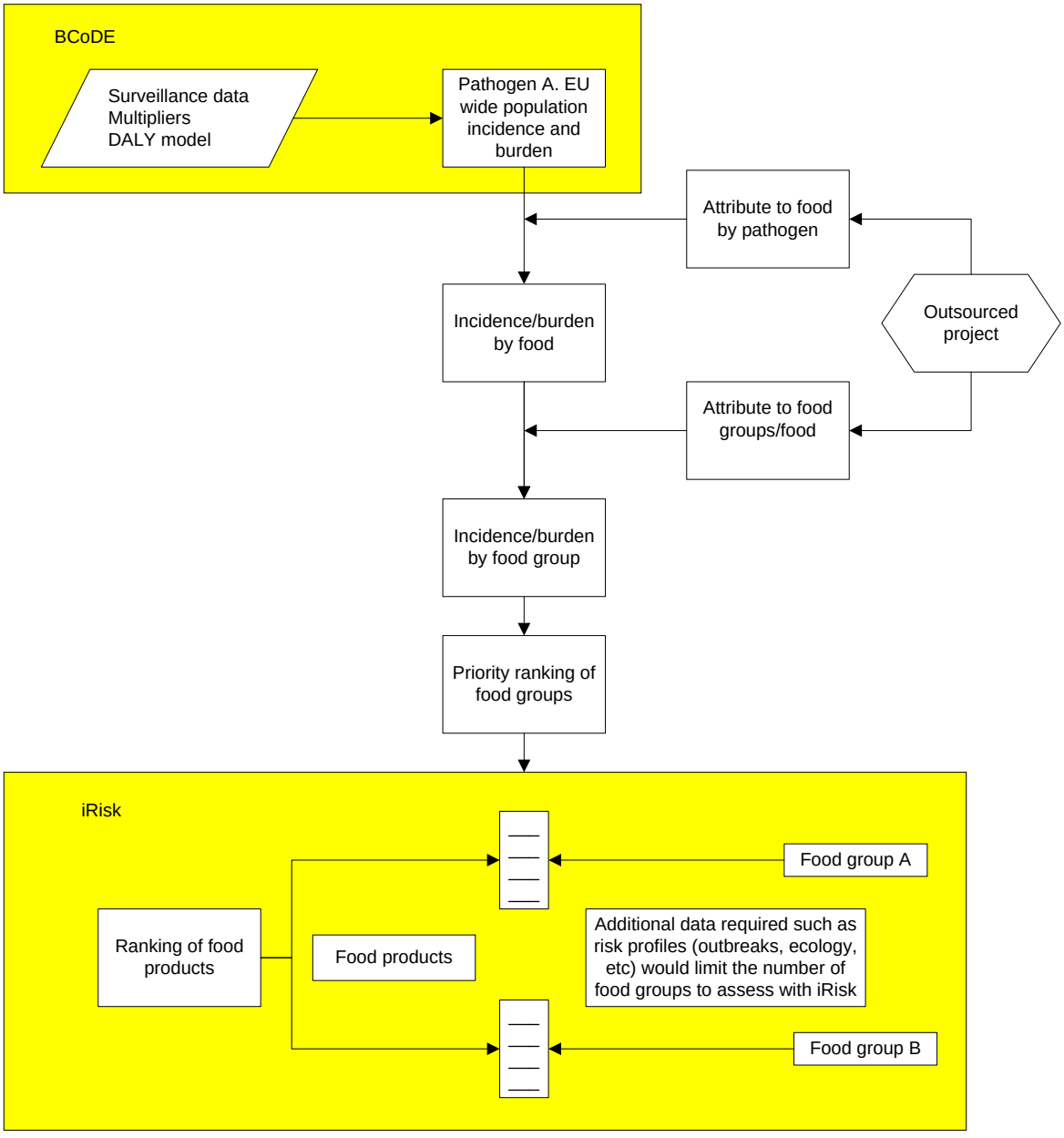
6.2. Top-down approach

BCoDE is a refined DALY calculator based on epidemiological data. When epidemiological data are available, the BCoDE toolkit is recommended for use by the BIOHAZ Panel. BCoDE was the only pure top-down risk ranking tool tested in this opinion and provides meaningful outputs such as DALYs, DALYs per case and DALYs per 100 000 cases. Risk ranking with BCoDE is based on a limited number of input parameters, namely the age group- and sex-specific number of cases, which reduces complexity of the tool while variability and uncertainty of all variables is taken into account using Monte Carlo simulation. As described in Section 2.8, it offers default values for the parameters, which can be modified by the user if specific data are available. The tool has an advanced user interface that allows effective data management and scenario analysis while outputs are presented in communication-friendly visualisations such as tables, bubble charts and bar charts. The current aim of BCoDE is to rank pathogens irrespective of transmission pathways. For its application in a food safety context, attribution of the proportion of cases transmitted by food and by specific food pathways (meat, dairy, produce, etc.) needs to be estimated separately.

6.3. Combining the bottom-up and top-down risk ranking approaches

Based on the experience of the EFSA BIOHAZ Panel, risk ranking of an increased number of pathogen/food combinations using a bottom-up quantitative approach is a laborious process and often difficult to complete within the usual time frame of the mandates. Thus, there is often a need to reduce the number of the food–pathogen pairs, focusing on those that present the greatest risk to public health. This opinion therefore proposes a combination of a bottom-up and top-down risk ranking approaches using the risk ranking tools FDA-iRISK and BCoDE, respectively. The combined approach starts with an initial priority ranking using the BCoDE tool, which limits the number of food–pathogen combinations based on the available epidemiological data. To complete this first step in a timely manner, an outsourced project may be needed for pathogen incidence attribution to foods and food categories at the EU level. In a second step, the number of pathogen/food pairs is further decreased, based on data and information in risk profiles including outbreaks, the microbial ecology, the growth ability of the pathogen, etc. In the later step, qualitative decision trees can be used as a tool for showing how decisions about classifying pathogens–food combinations into broad categories are made (e.g. inclusion/exclusion; high/low). In the last step, a quantitative bottom-up approach is applied to the remaining prioritised list of the pathogen/food combinations using the FDA-iRISK tool,

possibly combined with BCoDE for DALY estimation. The above steps are described in detail in the following diagrams for single pathogen–multiple food, multiple pathogens-single food and multiple pathogen–multiple food combinations (Figures 27, 28 and 29).



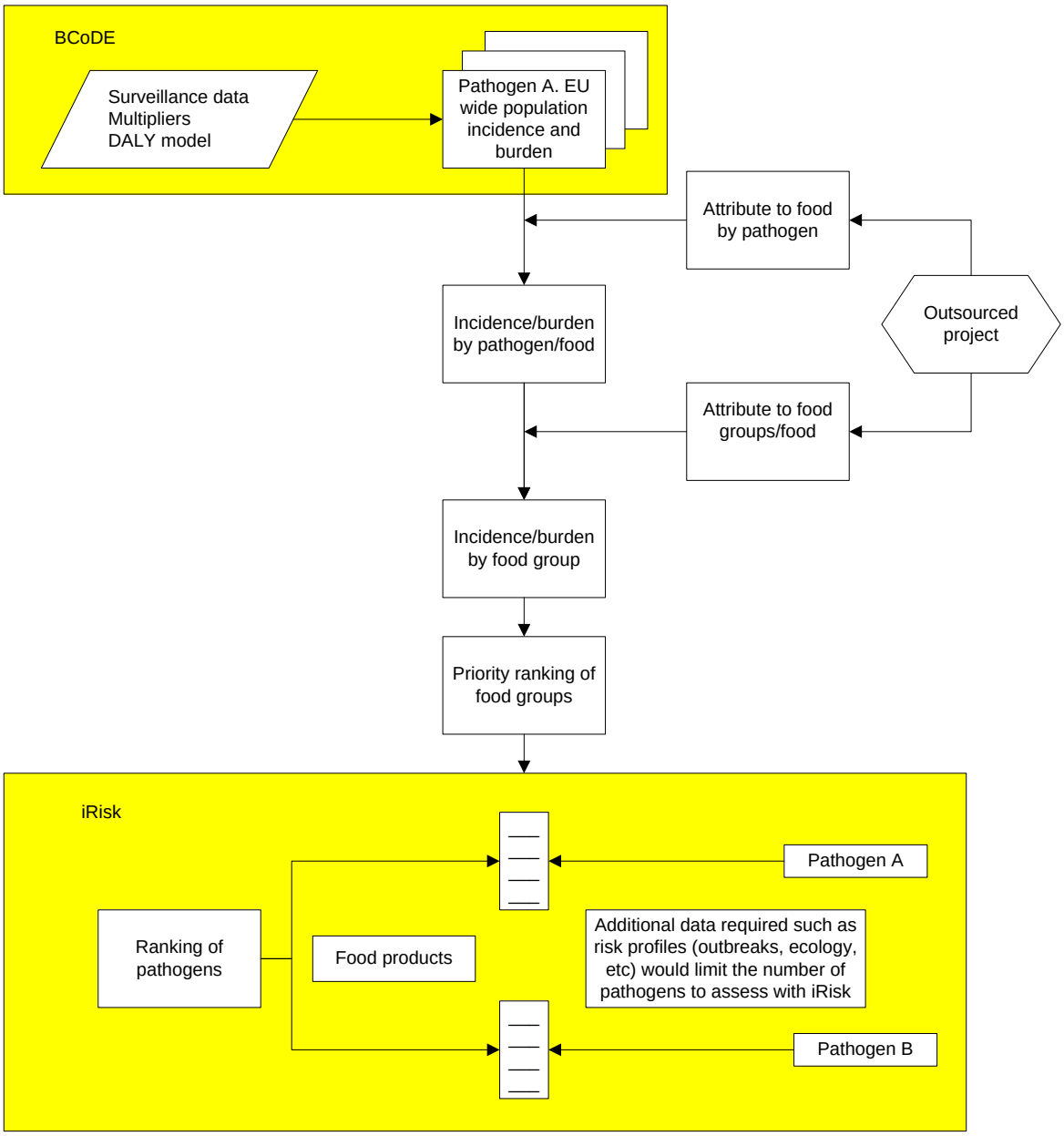
BCoDE: ECDC Burden of Communicable Diseases in Europe (toolkit); DALY: disability-adjusted life years.

Figure 27: Steps of the combined bottom-up and top-down risk ranking approach using the respective risk ranking tools FDA-iRISK and BCoDE for single pathogen–multiple food combinations

For a single pathogen/multiple foods question, the EU-wide population incidence and disease burden are estimated from surveillance data with appropriate multipliers and DALY models embedded in the BCoDE tool. Then, the proportion of disease cases and burden transmitted by food and by specific food groups are estimated using results from attribution studies. Such data are not yet available at the EU level, and a specific outsourced project could be considered to arrive at such estimates. A data-based approach would be preferable, but in the short-term expert elicitation may offer the most comprehensive approach. The study should conform to the EFSA Guidance on Expert Knowledge Elicitation (EFSA, 2014) and could be based on the protocol developed for the WHO by the

Foodborne Disease Burden Reference Group, which is described as one of the case studies in the EFSA Guideline.

Attribution data will assign the disease incidence and burden to broad food groups (to be further defined in the outsourced project), and may be the basis for not considering some food–pathogen pairs further. Risk profiles can be used to further limit the number of combinations to be quantitatively ranked using the FDA-iRISK tool, which will provide the final risk ranking output.

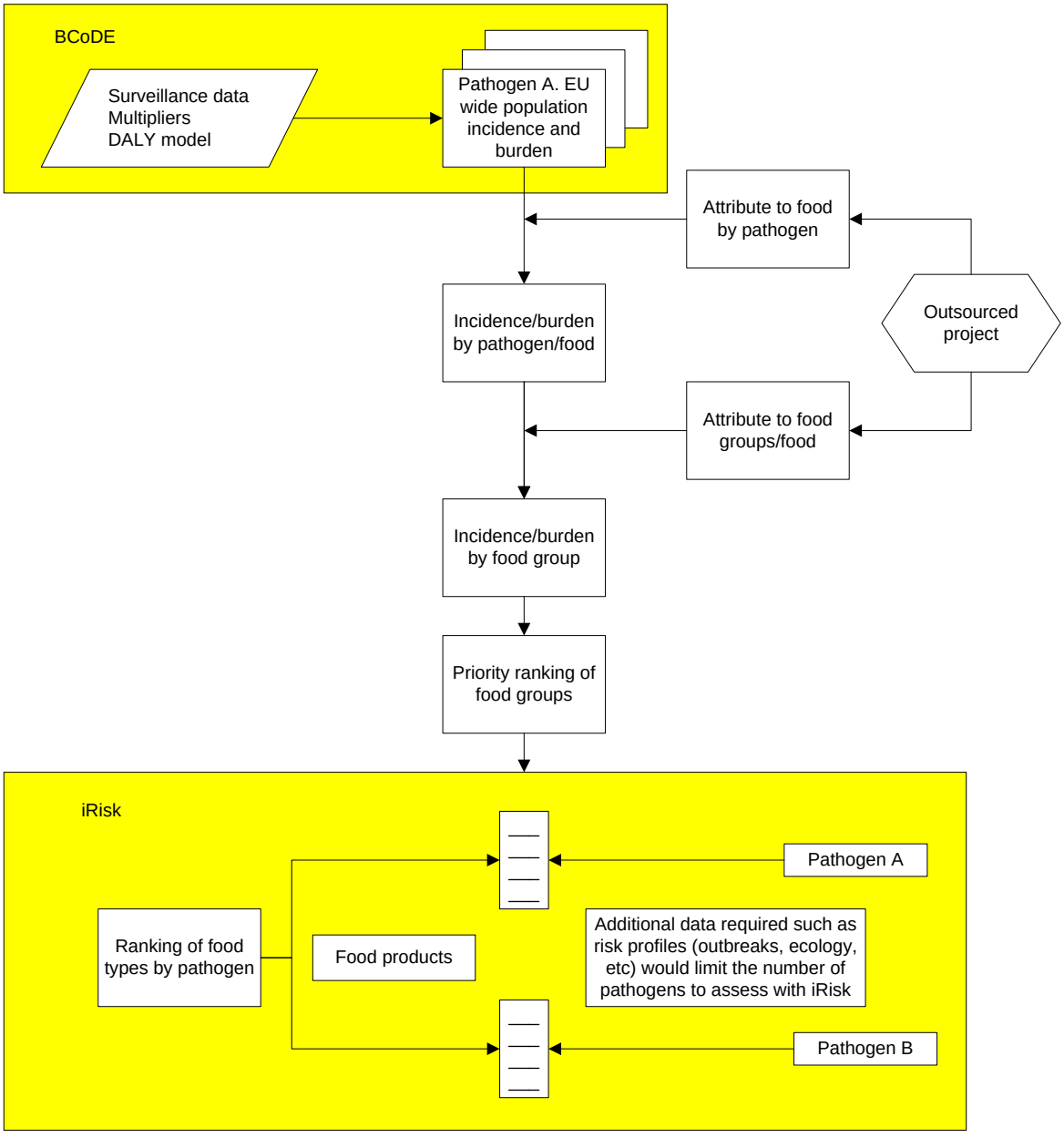


BCoDE: ECDC Burden of Communicable Diseases in Europe (toolkit); DALY: disability-adjusted life years.

Figure 28: Steps of the combined bottom-up and top-down risk ranking approach using the respective risk ranking tools FDA-iRISK and BCoDE for multiple pathogens–single food combinations

For ranking multiple pathogens in one food, the necessary epidemiological information needed as input for the BCoDE model and attribution data increase with the number of pathogens to be ranked. To improve efficiency, a proactive approach is suggested in which BCoDE models for key pathogens are developed as defaults by EFSA in collaboration with ECDC, while the outsourced attribution

project included the same pathogens. When this information is available, development of risk profiles can effectively be limited to pathogen/food pairs in the highest ranked combinations. As above, this would limit the number of combinations for quantitative ranking using FDA-iRISK.



BCoDE: ECDC Burden of Communicable Diseases in Europe (toolkit); DALY: disability-adjusted life years.

Figure 29: Steps of the combined bottom-up and top-down risk ranking approach using the respective risk ranking tools FDA-iRISK and BCoDE for multiple pathogen–multiple food combinations

If data and models needed for ranking single pathogens in multiple foods and multiple pathogens in single foods are available, the ranking of multiple pathogens in multiple foods is possible by following a combination of both approaches, although the workload to develop risk profiles and the number of pathogen/food pairs to evaluate using FDA-iRISK may increase considerably.

Any of the approaches described previously require appropriate documentation of data and the models used. There should be a clear and structured way of documenting all decisions made throughout a risk ranking exercise.

6.4. Development of new tool for BIOHAZ

The evaluation of the available tools showed a number of weak points and problems that may affect the risk ranking output. A main finding is that none of the available tools in their current form can describe uncertainty in risk ranking. Considering the importance of uncertainty in risk assessment showed in this opinion, as well as the other problems identified in the tested tools, this opinion also focused on the development of a new prototype risk ranking tool applicable to the whole chain, from farm to fork, for the EFSA BIOHAZ Panel.

The prototype risk ranking tool was developed as a RExcel application which is an add-in for Microsoft Excel and allows access to the statistics package R within the Excel environment. RExcel includes features such as data transfer (matrices and data frames) between R and Excel in both directions, running R code directly from Excel, writing macros calling R to perform calculations without exposing R to the user and calling R functions directly from cell formulas, using Excel's auto-update mechanism to trigger recalculation by R.

The prototype tool allows for a two-dimensional Monte Carlo simulation, providing the user with the ability to take into account both variability and uncertainty and also separate them. The prototype is based on a structure that takes into account all the factors affecting the risk and follows the risk assessment paradigm respecting the laws of probability and calculus. The tool can provide both deterministic and stochastic outputs for risk ranking using single values or distributions for the input parameters, respectively. In the stochastic application, the user can run the tool with variability only or with both variability and uncertainty.

The structure of the prototype consists of four main sheets for the (1) input parameters, (2) growth parameters, (3) uncertainty inputs and (4) outputs as presented in Figures 30 to 33.

The development of an advanced user interface for this prototype in the future which will allow an easy data management and scenario analysis can lead to an effective risk ranking tool for the EFSA BIOHAZ Panel.

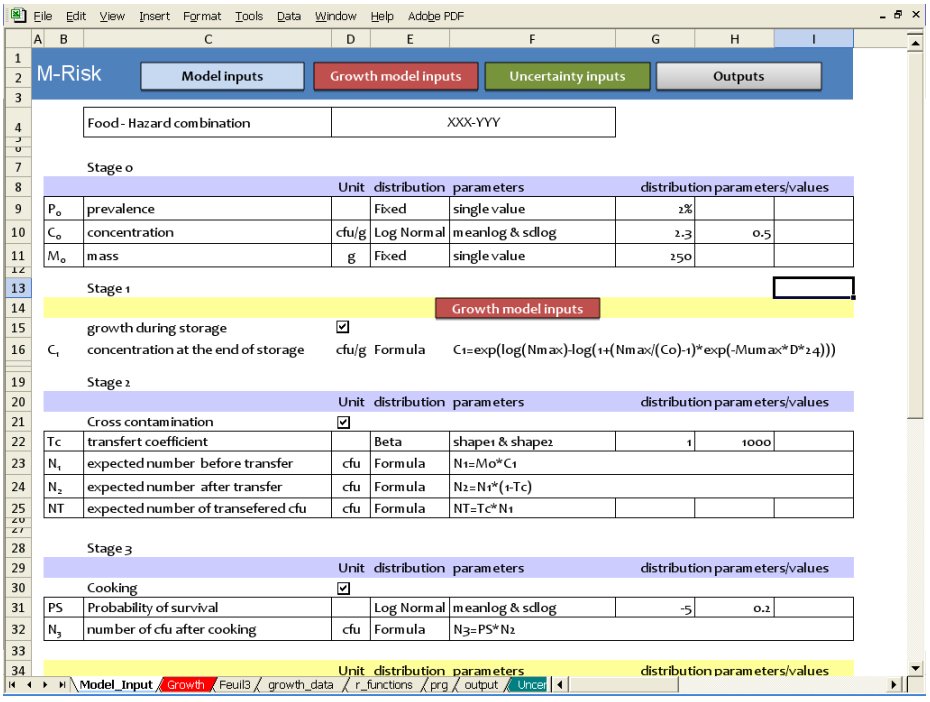


Figure 30: Input parameters sheet of the prototype risk ranking tool

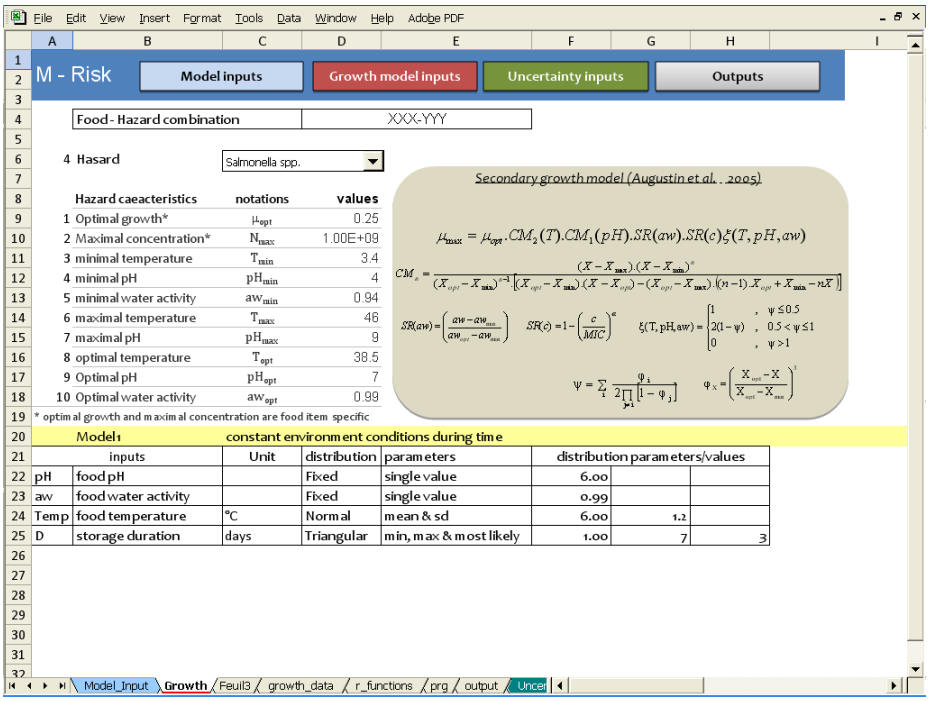


Figure 31: Growth parameters sheet of the prototype risk ranking tool

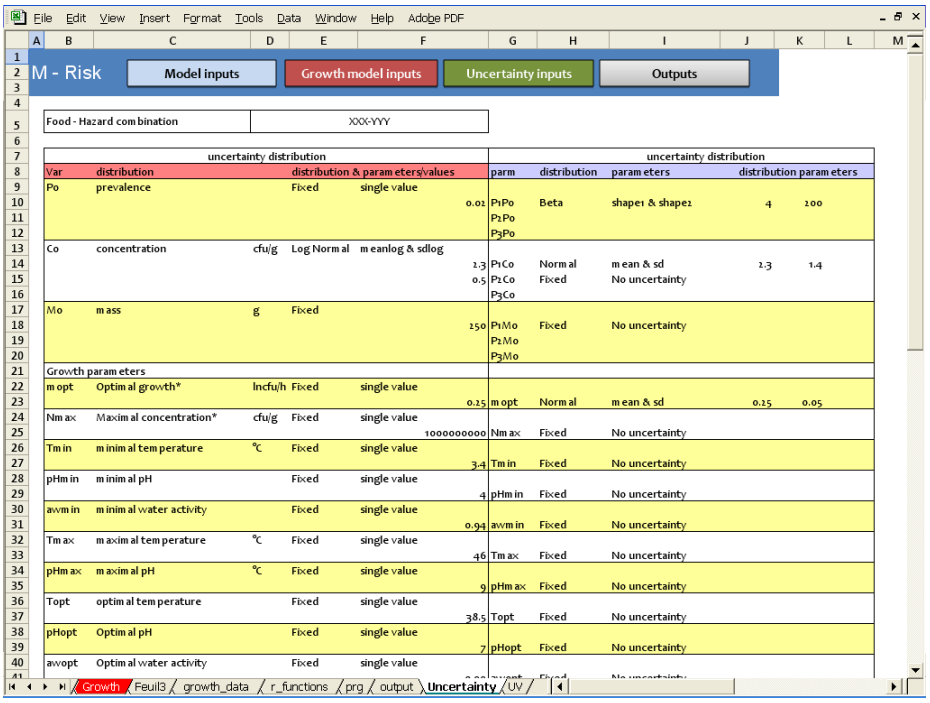


Figure 32: Uncertainty inputs sheet of the prototype risk ranking tool

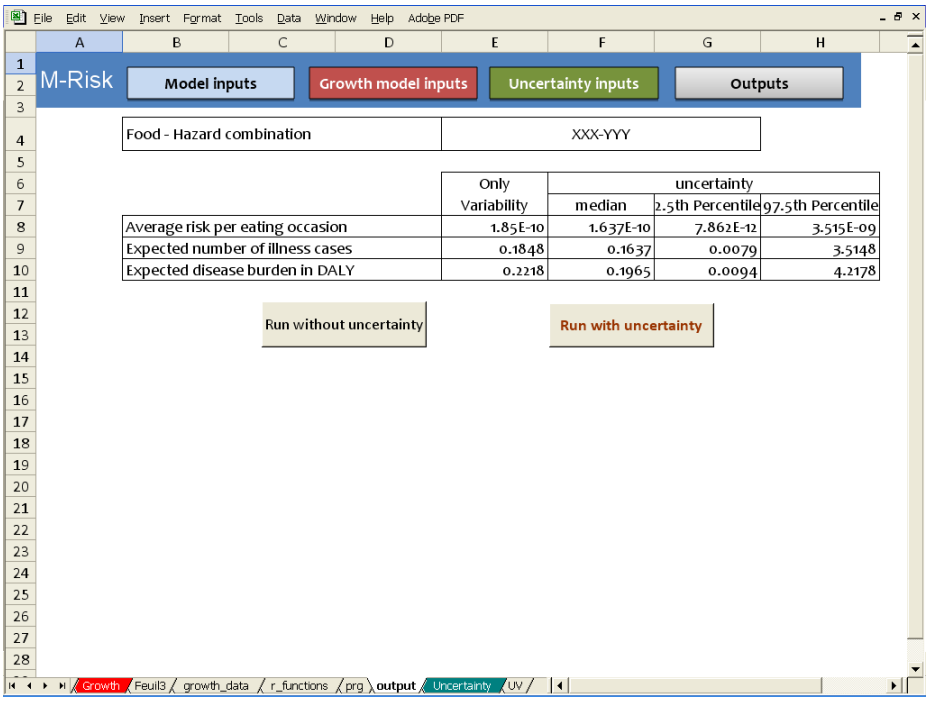


Figure 33: Outputs sheet of the prototype risk ranking tool

6.5. Networking of risk ranking tools with other available supporting tools, databases and information sources

The experience of EFSA BIOHAZ Panel has shown that the collection of data required for the development of risk ranking models, currently performed via literature review, is among the most laborious and time-consuming process in risk ranking exercises. In addition, literature review usually creates problems in the documentation and the transparency of the risk ranking models. As an alternative, the risk ranking tools proposed for the EFSA BIOHAZ Panel's use could be connected with other available tools, databases and information sources which can support the development of risk ranking models by exchanging models, data and information.

6.5.1. Description available supporting tools, databases and information sources

Apart from the risk ranking tools presented in this opinion, the last two decades of research has also focused on the systematic collection, description and modelling of food safety data, information and knowledge and their incorporation to user-friendly software and databases. A description of the above material, categorised based on the nature of support that can be provided to risk ranking is given below.

6.5.1.1. Predictive microbiology software tools

Predictive microbiology (PM) is an important part of risk assessment/ranking for the evaluation of the effect of food processing, storage and handling on the behaviour of pathogenic microorganisms. PM has established itself as a scientific discipline that uses mathematical equations to summarise and make readily available quantitative information on the microbial responses in various foods under different conditions (McMeekin et al., 2008). Development of models to predict survival, growth or inactivation of microorganisms in foods has been a most active research area within food microbiology during the last 25 years (Ross and Dalgaard, 2004). A considerable number of predictive microbiology software tools are today available to predict growth, survival and inactivation of microorganisms in foods and are described in the sections below.

Pathogen modelling program (PMP)

PMP is available free of charge (http://portal.arserrc.gov/) and, with more than 5 000 downloads per year, it is probably the most widely used predictive microbiology application software. PMP has been available for close to 20 years and it is regularly updated and expanded. The present version includes more than 40 models for different bacterial pathogens. The software allows growth or inactivation of pathogens to be predicted for different combinations of constant temperature, pH, NaCl/a_w and, in some cases, other conditions such as organic acid type and concentration, atmosphere or nitrate. In addition, PMP includes models that predict the effect of cooling temperature profiles on growth of *Clostridium botulinum* and *Cl. perfringens* after cooking. Predictions can be exported and the software contains references to studies from which the models were developed. In 2007, PMP was integrated with the Predictive Microbiology Information Portal (http://portal.errc.ars.usda.gov/).

ComBase (combined database on predictive microbiology information)

ComBase (www.combase.cc) is a web-based resource for Quantitative and Predictive Food Microbiology. Its main components are: a database of observed microbial responses to a variety of food-related environments and a collection of relevant predictive models. ComBase is managed by the ComBase Consortium consisting of the Institute of Food Research (IFR) in the United Kingdom, the US Department of Agriculture Agricultural Research Service (USDA-ARS) in the USA and the University of Tasmania Food Safety Centre (FSC) in Australia. The ComBase Predictive models are a collection of software tools based on ComBase data to predict the growth or inactivation of microorganisms. Currently available predictive tools include the following online applications:

- *ComBase Predictor*, a set of 23 growth models and six thermal death models for predicting the response of many important food-borne pathogenic and spoilage microorganisms to key environmental factors. An Excel version of this web application can also be found in the ComBase Excel Demo provided in the website
- *Perfringens Predictor*, an application specifically designed for predicting the growth of *Cl. perfringens* during the cooling of meats. An Excel Add-In version of the program can also be found in the downloads section of this website

Sym'previus

Sym'previus (www.symprevius.org) is an extensive French decision support system that includes (1) a database with growth and inactivation responses of microorganisms in foods and (2) predictive models for growth and inactivation of pathogenic bacteria and some spoilage microorganisms. Information from Sym'previus is available on a commercial basis through contact centres as indicated on the homepage cited above.

Seafood spoilage and safety predictor (SSSP)

The Seafood Spoilage and Safety Predictor (SSSP) software has been developed by Danish Technical University (http://sssp.dtuaqua.dk/HTML_Pages/Help/English/Index.htm) to facilitate the practical use of mathematical models to predict shelf life as well as growth of spoilage and pathogenic bacteria in seafood. SSSP v. 3.1 from August 2009 includes: four product-specific relative rate of spoilage (RRS) models, three generic RRS models, four product-specific microbial spoilage models, a generic model to predict microbial growth and shelf life, modules to compare predictions from SSSP with users own data of shelf life or growth of bacteria, models to predict growth and histamine formation by *Morganella psychrotolerans* and *M. morganii*, growth and growth boundary model for *L. monocytogenes* and a model to predict the simultaneous growth of *L. monocytogenes* and lactic acid bacteria in lightly preserved seafood.

GroPIN modelling database

GroPIN is an integrated tertiary model developed by Agricultural University of Athens (AUA) using Visual Basic for Applications (http://www.aua.gr/psomas/gropin/). The application may serve as a

user-friendly and highly transparent predictive modelling database for kinetic (growth or inactivation) and probabilistic models. The current version of GroPIN has a total of 490 published models for the behaviour of 22 pathogens and 50 spoilage organisms, including spoilage and mycotoxigenic fungi, bacteria and yeasts in various foods of plant (e.g. fresh-cut salads, deli salads, berries, juices, etc.) or animal origin (meat and meat products, dairy products). The impact on microbial behaviour of a variety of critical and commonly encountered intrinsic (preservatives, organic acids in total or undissociated/dissociated form, salt, a_w, nitrates, etc.) and extrinsic (temperature, CO_2, pressure, anaerobic conditions) factors is accounted for by the models registered in GroPIN up to date. The microbial responses modelled (i.e. dependent variables) include the maximum specific growth rate, the death rate, the lag phase duration, maximum population density, time to X-log reduction/growth, D-values and the probability of growth. The spirit of the software stems from similar initiatives, such as Sym'previus and COMBASE modelling toolbox. The major innovative features of this software in relation to the state-of-the art are the user-friendliness, the updatable character by the user, the simplicity and functionality (including interactive options) of outputs and the inclusion of all major predictive modelling classes.

Refrigeration index calculator

Refrigeration index (RI) calculator was developed by Meat & Livestock Australia Limited (http://www.foodsafetycentre.com.au/refrigerationindex.php). It predicts the expected growth of *E. coli* on meat from temperature and other data. The model has values for pH, a_w and lactate concentration which in addition to temperature, all affect the growth rate of *E. coli*. The current RI model allows for the user to enter data on temperatures of the product over time. The other parameters are set by choosing the type of product.

Opti-Form@ *Listeria* control model 2007 (PURAC)

This software predicts the effect of organic acids, temperature, pH and moisture on growth of *L. monocytogenes* in meat products. The software can be requested from the PURAC company (http://www.purac.com/purac_com/d9ed26800a03c246d4e0ff0f6b74dc1b.php).

Shelf stability predictor

The software has been developed by the Center for Meat Process Validation at the University of Wisconsin (Madison, USA) (http://meathaccp.wisc.edu/ST_calc.html) and provides a set of models for predicting the growth of *L. monocytogenes* and *S. aureus* on RTE meat products as a function of pH and water activity.

THERM (temperature history evaluation for raw meat)

Developed by the Center for Meat Process Validation at the University of Wisconsin (Madison, USA) (http://meathaccp.wisc.edu/). THERM is an online tool designed for evaluating the safety of meat or poultry at temperatures between 50 °F and 115 °F (10 °C to 46 °C)

Process lethality determination spreadsheet

Developed by AMI Foundation, USA, (http://www.amif.org/ht/d/sp/i/26870/pid/26870) this tool provides processors with a science-based validation tool that can be used to demonstrate the effectiveness of a specific heat process to destroy a microorganism of concern. Specifically, the interactive model allows the user to input actual in-process data from a given cook cycle and determine if the process achieves the required log reduction for the microorganism of concern. The goal is to define or map the heating and cooling profile of the product by observing the temperature characteristics of the product during heating and cooling.

6.5.1.2. Databases that can provide input data for risk ranking

EFSA zoonoses, antimicrobial resistance and food-borne outbreaks

EFSA analyses data on zoonoses, antimicrobial resistance and food-borne outbreaks across the EU. Data are submitted annually by the MS. Zoonoses are infections and diseases that are transmissible between animals and humans. EFSA publishes, in collaboration with the ECDC, annual Community Summary Reports based on these data. ECDC provides for and analyses data on the zoonoses cases in humans. The latest report covers 18 zoonotic infections. Moreover, EFSA analyses the EU-wide baseline surveys on zoonotic agents, such as *Salmonella* and *Campylobacter*, in animal and food-populations and on antimicrobial resistance. These surveys are fully harmonised and therefore provide comparable values for all MS. Survey results are used to set EU reduction targets or to consider needs for specific actions at EU-level. The EFSA Zoonoses, antimicrobial resistance and food-borne outbreaks database can provide data for prevalence and concentration of pathogens in a bottom-up risk ranking approach as well as epidemiological data for top-down risk ranking.

FOSCOLLAB: a global platform for food safety data and information

FOSCOLLAB (http://www.who.int/foodsafety/foscollab/en/) is a new WHO platform for food safety professionals that enables users to: (1) access food safety data and information quickly, (2) maximise the utility of already existing sources and minimise duplication of effort, (3) integrate data and information coming from animal/agriculture, food and human health areas to improve global public health, (4) promote better generation of data, and (5) strengthen the underlying sources by promoting awareness and increased utilisation. By integrating multiple sources of reliable data, FOSCOLLAB helps overcome the challenges of accessing these key sources in a timely manner. It allows for better risk assessment and decision-making by food safety professionals and authorities.

The EFSA Comprehensive European Food Consumption Database

The Comprehensive Food Consumption Database is a source of information on food consumption across the EU. It contains detailed data for a number of EU countries. The database plays a key role in the evaluation of the risks related to possible hazards in food in the EU and allows estimates of consumers' exposure to such hazards, a fundamental step in EFSA's risk assessment work. Summary statistics from the database enable quick screening for chronic and acute exposure. In the database, dietary surveys and food consumption data for each country are divided by category. These include: age, from infants to adults aged 75 years or older; food group (nearly 160) and type of consumption, covering both regular and high consumption thus allowing calculations to be tailored to each category of consumer. The statistics on food consumption are reported in grams per day, as well as grams per day per kilogram of body weight.

Food Commodity Intake Database (FCID) (http://fcid.foodrisk.org/)

The Food Commodity Intake Database (FCID) was developed by US EPA's Office of Pesticide Programs (OPP) to improve the utility of the food consumption survey for dietary exposure assessment. FCID 2003–2008 translates food consumption as reported eaten in What We Eat in America (WWEIA) (1999–2008 survey cycles) and Continuing Survey of Food Intakes by Individuals (CSFII) (1994–1996/1998) surveys into consumption of US EPA-defined food commodities. Such food commodity intakes are expressed as grams of food commodity consumed per kilogram of body weight per day for over 500 commodities derived from more than 6 000 different foods and beverages reported in the two surveys. WWEIA-FCID 2003–2008 is intended to complement the CSFII and National Health and Nutrition Examination Survey (NHANES)/WWEIA databases in that it provides estimates of food consumption expressed as food commodities as opposed to foods per se (i.e. "as eaten") which can in some exposure and other situations be of more utility. The database also includes WWEIA 2003–2008 food consumption and demographic data that are available through CDC's National Center for Health Statistics. FCID can provide risk ranking models with consumption data.

The European Surveillance System (TESSy)

TESSy (http://www.ecdc.europa.eu/en/activities/surveillance/tessy/Pages/TESSy.aspx) is a highly flexible metadata-driven system for collection, validation, cleaning, analysis and dissemination of data. Its key aims are data analysis and production of outputs for public health action. All EU MS (28) and EEA countries (3) report their available data on communicable diseases (49), as described in Decision No 2119/98/EC, to the system. Apart from routine surveillance, TESSy has replaced data collection systems in place for the Dedicated Surveillance Networks (DSNs) to provide experts with a one-stop shop for EU surveillance data. Prior to May 2005, when ECDC was established, there were 17 DSNs that collected data on a variety of diseases. All MS submitted data individually to every DSN, using different file specifications. The TESSy database can be a source for epidemiological data for top-down risk ranking approaches.

FRISBEE

FRISBEE (http://frisbee-project.eu) is a Food Refrigeration Innovation for Cold Chain research IP European project. Within FRISBEE, the Cold Chain Database (hosted in the link http://www.frisbee-project.eu/coldchaindb.html/) has been built for data collection of temperature conditions throughout the food supply chain for different chilled and frozen food products. A systematic data collection for identification and evaluation of the weak links of the cold chain for different types of chilled and frozen products took place. Data from industry, cold chain parties (distributors, retailers) and consumer surveys, including all stages of the cold chain (from production to consumption) were collected. The Cold Chain Database has been constructed in order to develop a user-friendly online platform where collected data are retrievable and available to be used from candidate users. Registered Cold Chain Database users are able to retrieve specific time–temperature profiles using a multi-search criteria search engine. Stage/step of the cold chain, food storage temperature range, characterisation of food, food products, etc., are included in the available search criteria. At present, the Cold Chain Database consists of more than 11 500 time–temperature profiles and is being continuously updated with new data uploaded from an expanding network of contributors. In this database, the user can build a specific sequence of cold chain stages for specific food products based on user-defined search criteria. The Cold Chain Database can be used in combination with available predictive microbiology tools for the quantitative evaluation of pathogen's growth and/or survival during chilled storage and distribution.

6.5.1.3. Other risk assessment information sources

FoodRisk.org

FoodRisk.org is a repository for risk assessment data, information and tools. It is operated by Joint Institute for Food Safety and Applied Nutrition (JIFSAN) in collaboration with the Center for Food Safety and Applied Nutrition from US FDA (CFSAN/FDA) and the Food Safety and Inspection Services from USDA (FSIS/USDA). The aim of FoodRisk.org is to assist professionals involved with the many aspects of risk analysis as it pertains to the safety of our food. FoodRisk.org includes unique datasets, tutorials, tools and links to numerous sources of information. The goals identified for Foodrisk.org to date include consolidating risk analysis research data and methodology from public and proprietary sources, assisting coordination of research activities, identifying gaps in needed research and assisting the development of food safety risk assessment models. While initial emphasis was on microbial pathogens and their toxins, this is being expanded to other chemicals and toxins.

ICRA interactive online catalogue on risk assessment

The interactive online catalogue on risk assessment (ICRA) (http://icra.foodrisk.org/) is a repository of risk assessment models. ICRA was funded by the National Institute for Food and Agriculture (NIFA) of the USDA. It is a partnership between the National Institute for Public Health and Environment (RIVM) in the Netherlands, the National Food Institute (DTU Food) at the Technical University of Denmark, and JIFSAN at the University of Maryland. ICRA serves as a web tool offering a dynamic model catalogue for existing microbial risk assessments for risk assessors aiming to develop their own

models. ICRA allows users to compare and contrast models from the same pathogen and/or commodity. ICRA relies on contributions from risk assessors and modellers around the world to submit their models, populating the online catalogue.

6.5.2. Communication between risk ranking tools for EFSA BIOHAZ Panel use and other available supporting tools, databases and information sources

As described above, there is an increasing number of predictive microbiology software tools, databases and other information sources that could be used for extracting input data required for the development of risk ranking models using the tools proposed for EFSA. Of course, the use of above data requires validation for its applicability in each risk ranking exercise. After validation, the exploitation of the above supporting material could increase the transparency and reduce significantly the time required for performing a risk ranking. However, an important limitation for this is that up till now data and information exchange is difficult because of the lack of communication. The harmonisation of terms and concepts as well as the generation of information exchange formats are important issues in the field of food safety. The development of a communication language which will include a common file format for exchange of models/data/information as well as standards for description and documentation will allow an effective information exchange. Recently, the project "OpenML for Predictive Modelling in Food" was initiated (http://sourceforge.net/projects/microbialmodelingexchange/) as a community effort to establish an open information exchange data standard to facilitate free information exchange between different software tools developed within the community of predictive modelling in food. The extension of such initiations to risk assessment and the harmonisation of all related sources could lead to an effective toolbox of the risk assessment/ranking tools such as the FDA-iRISK and BCoDE combined with a supporting network of predictive microbiology tools and databases (Figure 34).

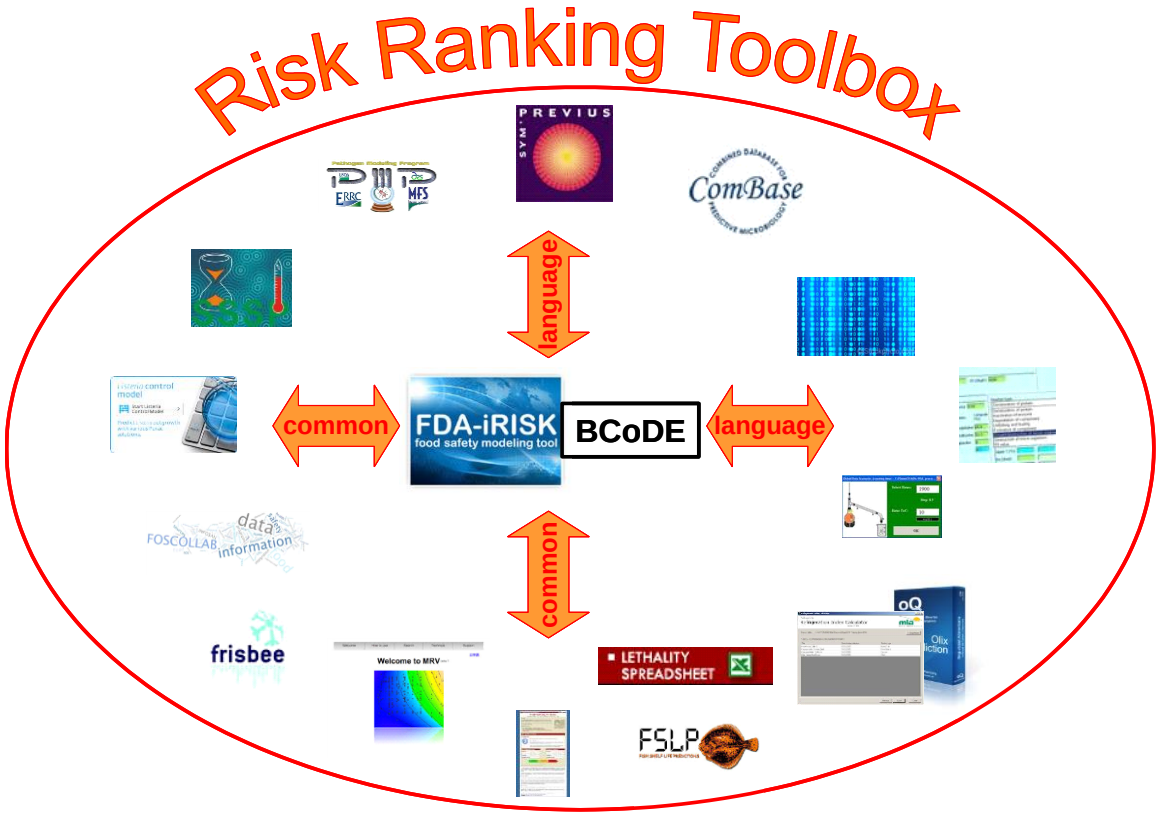


Figure 34: Representation of the risk assessment/ranking toolbox consisted of the FDA-iRISK/BCoDE tools and a supporting network of predictive microbiology tools and databases

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

ToR1. To evaluate the performance and data requirements of the available risk ranking tools

- Eight tools relevant to risk ranking applications of biological hazards in food were identified: decision trees; the United States Food and Drug Administration (US-FDA) risk ranking tool; the pathogen–produce pair attribution risk ranking tool (P^3ARRT); the EFSA food of non-animal origin risk ranking tool (EFoNAO-RRT); Risk Ranger; microHibro; swift quantitative microbiological risk assessment (sQMRA); FDA-iRISK and the European Centre for Disease Prevention and Control (ECDC) Burden of Communicable Diseases in Europe (BCoDE) toolkit.
- A detailed description of the tools, based on the conceptual risk ranking framework developed by the BIOHAZ Panel and their use in two risk ranking case studies showed clear differences among them related to the risk metrics, the ranking approach, the model type, the model variables and data integration method.
- Risk ranking tools have different data requirements, and empirical data requirements increase moving from qualitative to quantitative risk ranking approaches.
- Due to the differences in the tools, they provide different ranking results when applied to the case studies of single pathogen–multiple foods (*Listeria monocytogenes* in ready-to-eat (RTE) foods) and multiple pathogens in a single food (leafy greens).
- The selection of the risk metric was found to significantly affect the risk ranking because the metrics measures different things, for example probability of illness versus public health burden (disability-adjusted life years (DALYs)). Summary measures of public health such as DALYs integrate disease incidence, severity and mortality in a single number.
- Decision trees provide an arbitrary outcome and have very limited discrimination power for risk ranking. However, they have fewer data and time requirements and can be used to increase transparency when classifying risks into broad categories.
- Fully quantitative stochastic models are the most reliable for risk ranking. However, this approach needs a good characterisation of the input parameters.
- The use of deterministic models that ignore variability may result in risk ranking errors, which may be greater for the food–pathogen combinations with the highest risk, as shown in the example.
- In deterministic approaches, the selection of the point estimate used in the model can affect the risk ranking. Among different possible point estimates (arithmetic mean, median, 75^{th} and 90^{th} percentiles), the use of a high percentile provides, in general, ranking results which are most similar to a stochastic model. However, the performance of different point estimates in a ranking assessment will depend on the data input for the specific food–pathogen combinations involved.
- When using semi-quantitative models with ordinal scoring, the food–pathogen combinations are classified into broad sets of categories with little discrimination. There are considerable differences in risk ranking compared with a quantitative stochastic model. The ordinal scoring approaches provide ranking with more errors than the deterministic approaches.
- Among the quantitative tools that use a bottom-up approach for risk ranking, FDA-iRISK has been identified as the most appropriate for the needs of EFSA BIOHAZ Panel. FDA-iRISK is a technically sound, quantitative tool providing meaningful risk metrics, allowing effective data management and scenario analysis.
- The evaluation of the FDA-iRISK identified some limitations including the omission of a maximum population density and the lack of uncertainty assessment. A new version of the FDA-iRISK addressing most, if not all, of these issues is expected to be available early in 2015.

- BCoDE is a flexible, detailed and user-friendly DALY calculator that can be used as a top-down tool based on epidemiological data to rank pathogens. It is possible to generate additional disease models or scenarios according to the foods that are evaluated and data that are available.
- The performance of the risk ranking tools selected was evaluated from a statistical/theoretical perspective. Their implementation in practice may be constrained by limitations in data, time and resources.

ToR2. To investigate methodologies for introducing uncertainty and variability in the risk ranking models

- Uncertainty in risk assessment and risk ranking has been defined by the EFSA as "all types of limitations in knowledge, at the time it is collected".
- Uncertainty may arise from several factors in the risk assessment/ranking and includes technical (inexactness), methodological (unreliability), epistemological (ignorance) and societal (limited social robustness) aspects.
- Uncertainty in risk ranking needs to be carefully addressed and communicated to decision makers and stakeholders as one of the outcomes of the risk ranking process.
- Different typologies of uncertainty are available and provide a framework to identify and characterise all sources of uncertainty in a risk assessment/ranking model, and to identify how to evaluate them on their own scale and their impact on the outcomes of the risk assessment/ranking.
- The NUSAP (numeral, unit, spread, assessment and pedigree) system aims to characterise and prioritise sources of uncertainty in a risk assessment/ranking model. NUSAP is a generic method that can be applied to all types of models and provides standardised scales for description of uncertainty in various dimensions.
- NUSAP uses expert judgement to evaluate the impact of uncertainty in individual model factors on the outcome of the assessment, leading to a prioritisation of factors for further work (e.g. sensitivity and scenario analysis, or stochastic modelling).
- The combination of uncertainty typology and NUSAP helped to systematically identify and evaluate the uncertainty sources related to model outcomes and to assess their impact on the end results in a case study, using EFoNAO-RRT.
- Applying the NUSAP method requires training of the experts involved to overcome ambiguity of language in the pedigree scales.
- Uncertainty in rank orders cannot be formally quantified using qualitative or semi-quantitative ranking methods even though these are often applied in situations where data are limited.
- Uncertainty and variability in risk ranking can be represented by means of probability distributions, for example using two-dimensional Monte Carlo simulations. However, probabilistic representation is difficult when sufficient data are not available for statistical analysis. Expert elicitation procedures to incorporate diffuse information into the corresponding probability distributions may be adopted.

ToR3. To design and develop a risk ranking toolbox for the EFSA BIOHAZ Panel

- BCoDE and FDA-iRISK can be the basis of a risk ranking toolbox for use by the BIOHAZ Panel, which can be applied based on a "fit for purpose" approach.
- The validity and utility of the tools can vary depending on, for example the scope of the risk question in terms of the resolution needed, resource constraints and the availability of data. Consequently, tiered or step-wise approaches can be useful.

- When sufficiently accurate data are available for a fully bottom-up quantitative model and a limited number of food–pathogen combinations are to be ranked, FDA-iRISK can be used. Output from FDA-iRISK can be combined with BCoDE for a more effective calculation of DALYs.
- When surveillance epidemiological data are available, the BCoDE toolkit is appropriate for use by the BIOHAZ Panel for a top-down risk ranking approach. In this case, attribution of specific transmission pathways is needed as input for BCoDE. For this purpose, an EU-wide attribution study is needed, for example by expert elicitation.
- When many pathogen/food combinations are to be ranked, the application of a combined bottom-up and top-down risk ranking approach using the risk ranking tools FDA-iRISK and BCoDE, respectively, is more appropriate. The combined approach includes an initial priority ranking using the BCoDE tool, which limits the number of pathogens based on available epidemiological data. In a next step, the number of food–pathogen combinations is further decreased based on data and information of their risk profiles. In the last step, a quantitative bottom-up approach is applied for the remaining food–pathogen combinations using the FDA-iRISK tool.
- The evaluation of the available tools showed that none of them in their current form takes into account uncertainty in risk ranking. Considering the importance of uncertainty, a new prototype risk ranking tool for the EFSA BIOHAZ Panel was developed as a RExcel application. The prototype tool allows for a two-dimensional Monte Carlo simulation providing the user with the ability to take into account and separate variability and uncertainty. Future development of this prototype is needed before it can be used as an effective risk ranking tool for the EFSA BIOHAZ Panel. Necessary developments include a better user interface that will allow easier data management and scenario analyses.
- The risk ranking tools proposed for EFSA BIOHAZ Panel in combination with a network of available predictive microbiology tools, databases and information sources can form a risk ranking toolbox. This toolbox will support the timely and transparent development of risk ranking by allowing access to models, data and information.

RECOMMENDATIONS

- Risk metrics used in risk ranking should have a meaningful biological or epidemiological interpretation and have to be agreed with the risk managers before starting the risk ranking exercise.
- Decision trees should only be used as a tool for showing how decisions about classifying pathogens–food combinations into broad categories are made (e.g. inclusion/exclusion; high/low).
- Quantitative risk ranking models respecting the rules of probability calculation and describing correctly the main biological phenomena that determine the risk are preferred over semi-quantitative models with ordinal scoring.
- Quantitative risk ranking models should preferably include variability. If this is not possible, deterministic models may be used, where more than one point estimates (e.g. arithmetic mean and a higher percentile) should be used as part of sensitivity analysis to compare rankings.
- A framework encompassing uncertainty typology and evaluation (for example by NUSAP) should preferably be part of each risk ranking process to formalise discussions on uncertainty, considering practicality and feasibility aspects.
- In the absence of representative and accurate data describing the variability and uncertainty, expert knowledge elicitation should preferably be carried out to assess the uncertainty about the key input parameters (identified using sensitivity analysis or the NUSAP approach for example).
- When data and time constraints do not allow quantitative risk ranking, semi-quantitative models could be used. In this case, the limitations of these approaches linked to the selection and integration of the ordinal scores, as identified in this opinion, should be made explicit.

- A strategy should be developed to progressively adopt the proposed methods in future risk ranking opinions developed by the BIOHAZ Panel.

REFERENCES

Anderson M, Jaykus LA, Beaulieu S and S D, 2011a. Pathogen-produce pair attribution risk ranking tool to prioritize fresh produce commodity and pathogen combinations for further evaluation (P3ARRT). Food Control, 22, 1865-1872.

Anderson M, Jaykus LA, Beaulieu S and Dennis S, 2011b. Pathogen-produce pair attribution risk ranking tool to prioritize fresh produce commodity and pathogen combinations for further evaluation (P(3)ARRT). Food Control, 22, 1865-1872.

Augustin JC, Zuliani V, Cornu M and Guillier L, 2005. Growth rate and growth probability of *Listeria monocytogenes* in dairy, meat and seafood products in suboptimal conditions. Journal of Applied Microbiology, 99, 1019-1042.

Baert L, Mattison K, Loisy-Hamon F, Harlow J, Martyres A, Lebeau B, Stals A, Van Coillie E, Herman L and Uyttendaele M, 2011. Review: norovirus prevalence in Belgian, Canadian and French fresh produce: a threat to human health? International Journal of Food Microbiology, 151, 261-269.

Baraldi P, Zio E and Compare M, 2009. A method for ranking components importance in presence of epistemic uncertainties. Journal of Loss Prevention in the Process Industries, 22, 582-592.

Bouwknegt M and Havelaar A, 2014. Uncertainty assessment using the NUSAP approach: a case study on the EFoNAO tool. EFSA supporting publication 2014:EN-663, 22 pp.

Carrasco E, Pérez-Rodríguez F, Valero A, García-Gimeno RM and Zurera G, 2010. Risk Assessment and Management of *Listeria Monocytogenes* in Ready-to-Eat Lettuce Salads. Comprehensive Reviews in Food Science and Food Safety, 9, 498-512.

CDC (Centers for Disease Control and Prevention), online. Foodborne Outbreak Online Database (FOOD). Available at: http://wwwn.cdc.gov/foodborneoutbreaks/

Chen Y, Dennis SB, Hartnett E, Paoli G, Pouillot R, Ruthman T and Wilson M, 2013. FDA-iRISK--a comparative risk assessment system for evaluating and ranking food-hazard pairs: case studies on microbial hazards. Journal of Food Protection, 76, 376-385.

CODEX (Codex Alimentarius Commission), 2007. Working principles for risk analysis for food safety for application by governments. Available at: www.codexalimentarius.net/input/download/standards/10751/CXG_062e.pdf

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

EFSA (European Food Safety Authority), 2014. Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment. EFSA Journal 2014;12(6):3734, 278 pp, doi:10.2903/j.efsa.2014.3734

EFSA and ECDC (European Food Safety Authority and the European Centre for Disease Prevention and Control), 2014. The European Union Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Food-borne Outbreaks in 2012. EFSA Journal 2014;12(2):3547, 312 pp, doi:10.2903/j.efsa.2014.3547

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), 2013. Scientific Opinion on the risk posed by pathogens in food of non-animal origin. Part 1 (outbreak data analysis and risk ranking of food/pathogen combinations). EFSA Journal 2013;11(1):3025, 138 pp, doi:10.2903/j.efsa.2013.3025

EFSA Panel on Biological Hazards (BIOHAZ), 2012a. Scientific Opinion on Public health risks represented by certain composite products containing food of animal origin. EFSA Journal 2012;10(5):2662, 132 pp, doi:10.2903/j.efsa.2012.2662

EFSA Panel on Biological Hazards (BIOHAZ), 2012b. Scientific Opinion on the development of a risk ranking framework on biological hazards. EFSA Journal 2012;10(6):2724, 88 pp, doi:10.2903/j.efsa.2012.2724

Evers E and Chardon J, 2012. A swift Quantitative Microbiological Risk Assessment (sQMRA) - tool: improved version. Proceedings of the 23rd International ICFMH Symposium FoodMicro 2012, Global Issues in Food Microbiology, 3-7 September 2012, Istanbul, Turkey, p. 10.

Evers E and Chardon J, 2013. A swift Quantitative Microbiological Risk Assessment (sQMRA) - tool: improved version. Proceedings of the Society for Risk Analysis, Risk Analysis for Better Policies, 2013 Annual Meeting, 8-11 December, Baltimore, Maryland, USA, p. 42.

Evers EG and Chardon JE, 2010. A swift Quantitative Microbiological Risk Assessment (sQMRA) tool. Food Control, 21, 319-330.

FAO/WHO (Food and Agriculture Organization of the United Nationals/World Health Organization), 2001. Codex Alimentarius - Food Hygiene - Basic Texts - Second Edition. Available at: http://www.fao.org/docrep/005/y1579e/y1579e00.htm.

FAO/WHO (Food and Agriculture Organization of the United Nationals/World Health Organization), 2002. Risk assessments of Salmonella in eggs and broiler chickens. Microbiological Risk Assessment Series 2. Available at: http://www.fao.org/docrep/005/Y4392E/Y4392E00.HTM

FAO/WHO (Food and Agriculture Organization of the United Nationals/World Health Organization), 2009. Risk assessment of *Campylobacter* spp. in broiler chickens: Technical Report. Microbiological Risk Assessment Series No 12. Geneva. 132 pp. Available at: http://www.fao.org/fileadmin/templates/agns/pdf/jemra/MRA_12.pdf.

Guillier L, Kabunda J-M, Denis J-B and Albert I, 2013. Elicitation for food microbial risk assessment: a probabilistic approach extending Risk Ranger proposal. Journal de la Société Française de Statistique, 154, 113-123.

Havelaar AH, Haagsma JA, Mangen MJJ, Kemmeren JM, Verhoef LPB, Vijgen SMC, Wilson M, Friesema IHM, Kortbeek LM, van Duynhoven Y and van Pelt W, 2012. Disease burden of foodborne pathogens in the Netherlands, 2009. International Journal of Food Microbiology, 156, 231-238.

Hoelzer K, Pouillot R, Gallagher D, Silverman MB, Kause J and Dennis S, 2012. Estimation of Listeria monocytogenes transfer coefficients and efficacy of bacterial removal through cleaning and sanitation. International Journal of Food Microbiology, 157, 267-277.

Horae CA, 1962. Quick Sort. Computer Journal, 5, 10-15.

ILSI (International Life Sciences Institute), 2010. Impact of microbial distributions of food safety. ILSI Europe report series. International Life Sciences Institute, Brussels. Available at: http://www.ilsi.org/Europe/Publications/Microbial%20Distribution%202010.pdf

Knol AB, Petersen AC, van der Sluijs JP and Lebret E, 2009. Dealing with uncertainties in environmental burden of disease assessment. Environmental Health, 8.

Koseki S and Isobe S, 2005a. Growth of *Listeria monocytogenes* on iceberg lettuce and solid media. International Journal of Food Microbiology, 101, 217-225.

Koseki S and Isobe S, 2005b. Prediction of pathogen growth on iceberg lettuce under real temperature history during distribution from farm to table. International Journal of Food Microbiology, 104, 239-248.

Marklinder IM, Lindblad M, Eriksson LM, Finnson AM and Lindqvist R, 2004. Home storage temperatures and consumer handling of refrigerated foods in Sweden. Journal of Food Protection, 67, 2570-2577.

McMeekin T, Bowman J, McQuestin O, Mellefont L, Ross T and Tamplin M, 2008. The future of predictive microbiology: Strategic research, innovative applications and great expectations. International Journal of Food Microbiology, 128, 2-9.

Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM and Tauxe RV, 1999. Food-related illness and death in the United States. Emerging Infectious Diseases, 5, 607-625.

Nauta MJ, 2005. Microbiological risk assessment models for partitioning and mixing during food handling. International Journal of Food Microbiology, 100, 311-322.

Nauta MJ, 2008. The modular process risk model (MPRM): A structured approach to food chain exposure assessment. Ed Schaffner DW. pp.

Nauta MJ, Jacobs-Reitsma WF and Havelaar AH, 2007. A risk assessment model for *Campylobacter* in broiler meat. Risk Analysis, 27, 845-861.

Petersen AC, Cath A, Hage M, Kunseler E and van der Sluijs JP, 2011. Post-Normal Science in Practice at the Netherlands Environmental Assessment Agency. Science Technology & Human Values, 36, 362-388.

Robert C and Casella G, 2004. Monte Carlo Statistical Methods. Ed DeVeaux R, Fienberg, Stephen E., Olkin, Ingram. Springer Science + Business Media New York, 649 pp.

Robertson LJ and Gjerde B, 2001. Occurrence of parasites on fruits and vegetables in Norway. Journal of Food Protection, 64, 1793-1798.

Ross T and Dalgaard P, 2004. Secondary models. In: Modeling microbial responses in food. Eds McKellar RC and Lu X, 63-150.

Ross T and Sumner J, 2002. A simple, spreadsheet-based, food safety risk assessment tool. International Journal of Food Microbiology, 77, 39-53.

Rosso L, Lobry JR and Flandrois JP, 1993. An unexpected correlation between cardinal temperatures of microbial growth highlighted by a new model. Journal of Theoretical Biology, 162, 447-463.

Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL and Griffin PM, 2011. Foodborne Illness Acquired in the United States-Major Pathogens. Emerging Infectious Diseases, 17, 7-15.

Strachan NJC, Doyle MP, Kasuga F, Rotariu O and Ogden ID, 2005. Dose response modelling of Escherichia coli O157 incorporating data from foodborne and environmental outbreaks. International Journal of Food Microbiology, 103, 35-47.

Teunis P, Van der Heijden O, Van der Giessen J and Havelaar A (RIVM), 1996. The dose-response relation in human volunteers for gastro-intestinal pathogens. National Institute of Public Health and the Environment, Bilthoven, 1996. Report no. 284550002.

Teunis PF, Chappell CL and Okhuysen PC, 2002. *Cryptosporidium* dose response studies: variation between isolates. Risk Analysis, 22, 175-183.

Teunis PF, Moe CL, Liu P, Miller SE, Lindesmith L, Baric RS, Le Pendu J and Calderon RL, 2008. Norwalk virus: how infectious is it? Journal of Medical Virology, 80, 1468-1476.

U.S. FDA (United States Food and Drug Administration), 2003. Quantitative assessment of relative risk to public health from foodborne *Listeria monocytogenes* among selected categories of ready-to-eat foods. Available at: http://www.fda.gov/downloads/food/foodscienceresearch/ucm197329.pdf

van der Sluijs JP, Craye M, Funtowicz S, Kloprogge P, Ravetz J and Risbey J, 2005. Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: the NUSAP system. Risk Analysis, 25, 481-492.

Wijnands LM, Delfgou-van Asch EH, Beerepoot-Mensink ME, van der Meij-Florijn A, Fitz-James I, van Leusden FM and Pielaat A, 2014. Prevalence and concentration of bacterial pathogens in raw produce and minimally processed packaged salads produced in and for the Netherlands. Journal of Food Protection, 77, 388-394.

Yilmaz E, Aslam JA and Robertson S, 2008. A New Rank Correlation Coefficient for Information Retrieval. SIGIR' 08, 587-594.

Appendix A. Additional information about risk ranking tools

FDA-iRISK

Graphical representation of how the inputs are integrated in the FDA-iRISK tool

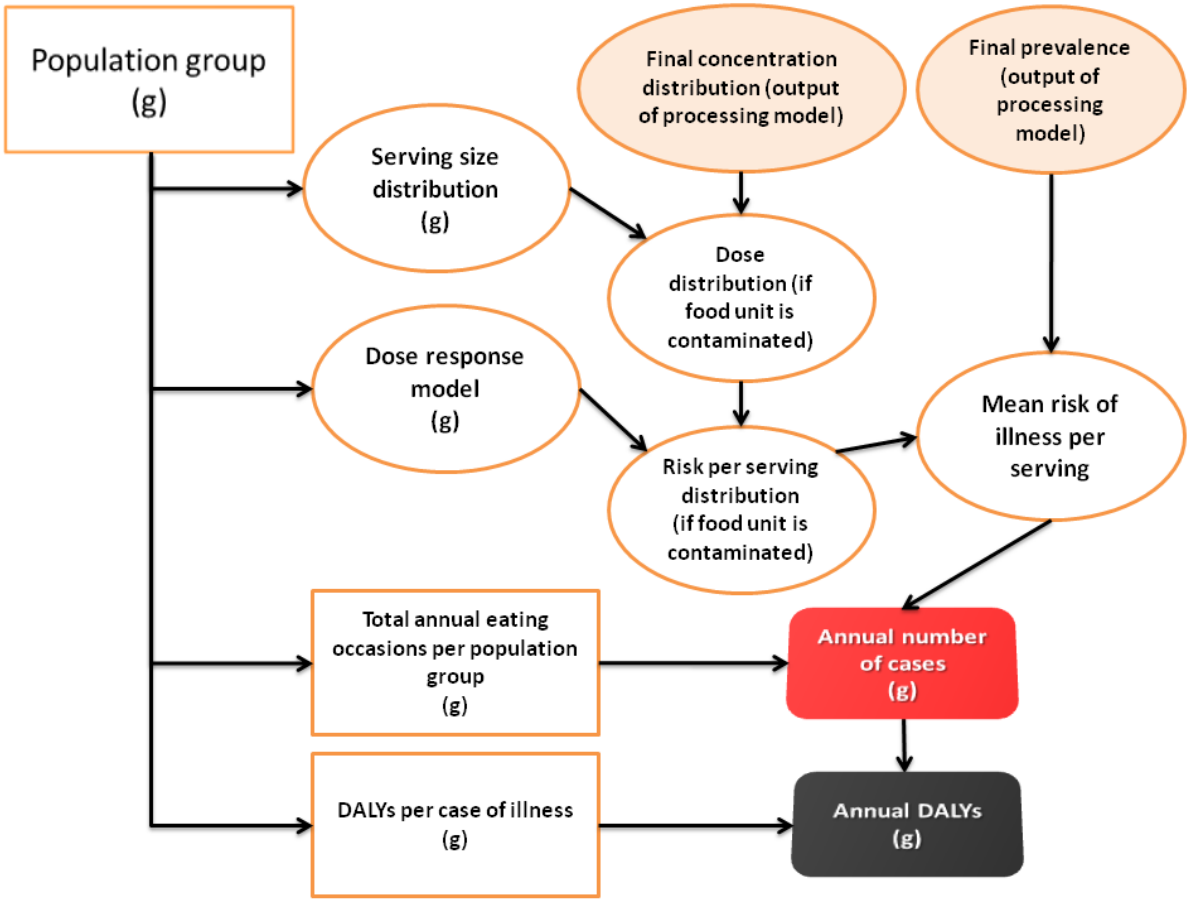
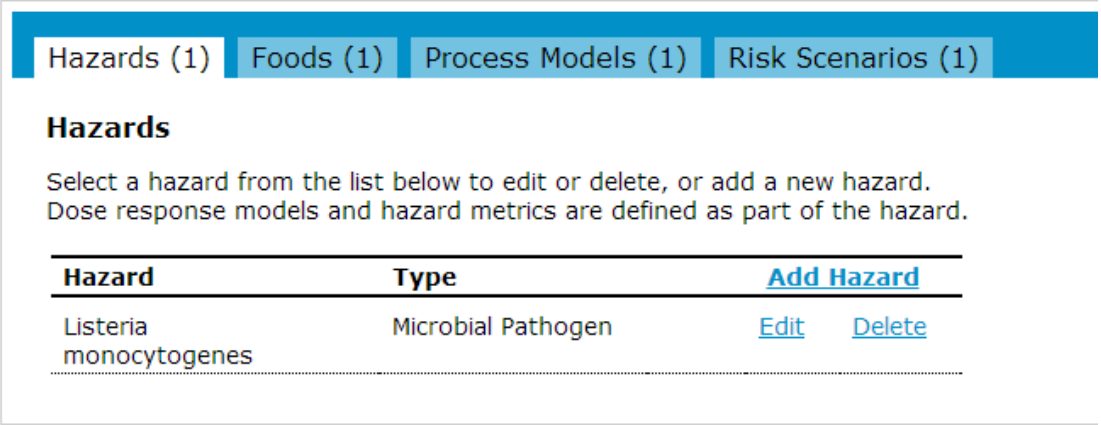


Figure 1: Integration of inputs in FDA-iRISK

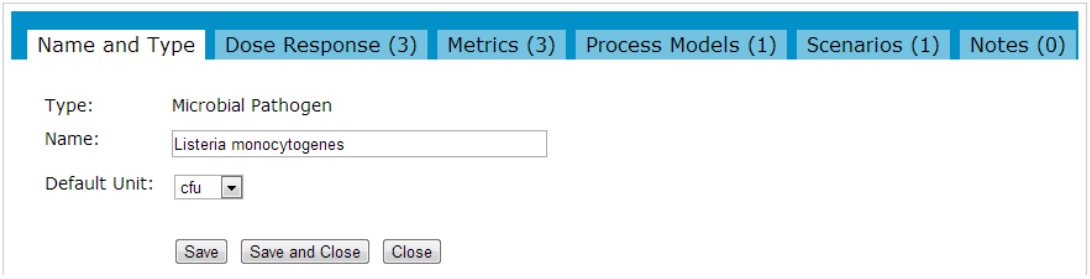
Examples of the FDA-iRISK interface

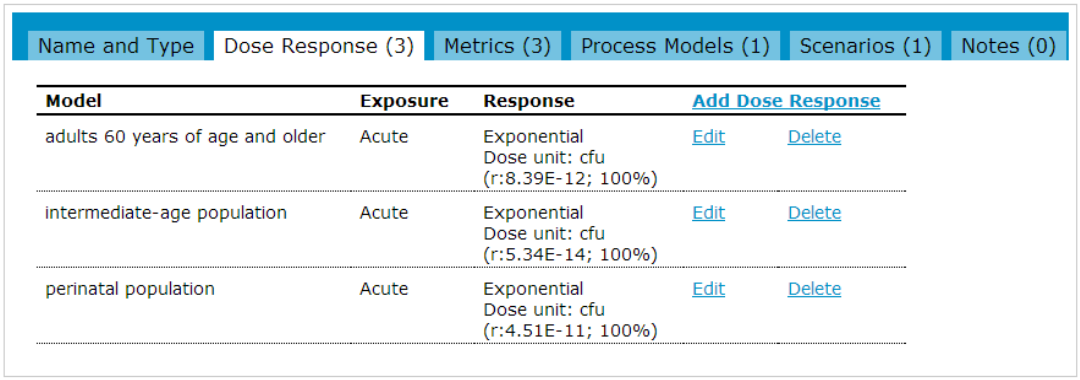


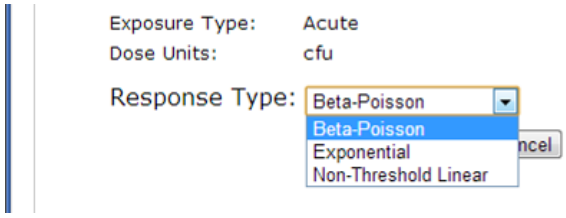
Edit Hazard

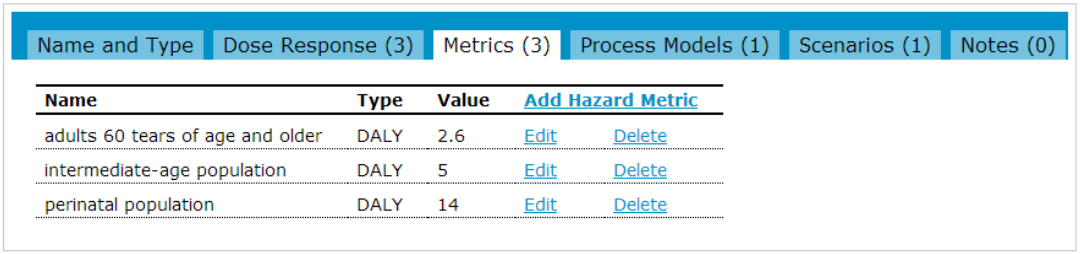
Use the "Name and Type" tab to edit the name and default units for the hazard. Use the remaining tabs to
elements to the hazard.

A hazard requires at least one dose response model and one hazard metric to be included in a computed ri

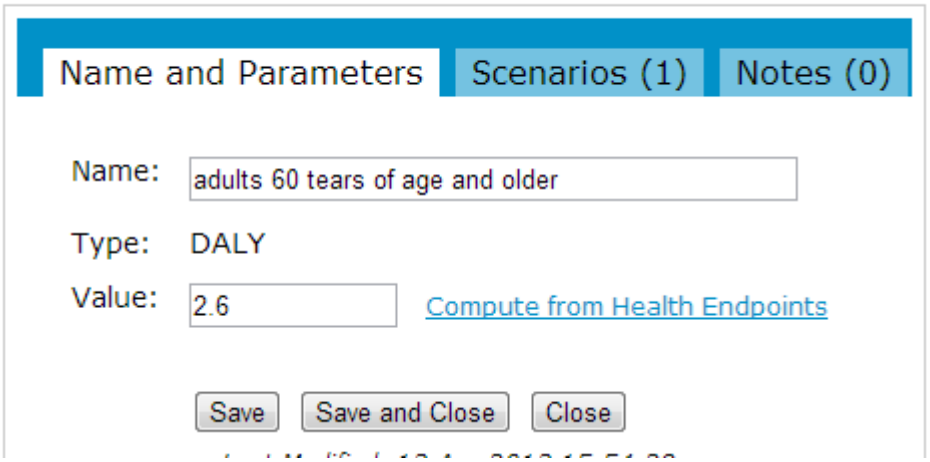








Quick Links: Hazards



Compute DALY

Use the Add button to add new health endpoints. Each endpoint must have a duration value in either years (Y) or days (D), a severity value ranging from 0 to 1, and a fraction of cases value (typically between 0 and 1).

The individual health endpoints are combined to compute the total DALY measure.

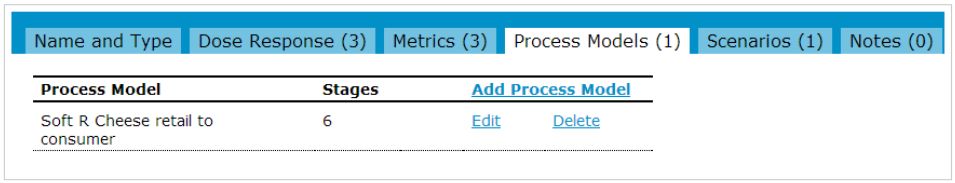
Health Endpoint	Duration	Unit	Severity	DALY	Fraction of Cases	Weighted DALY	
		Y					Add
Totals:					0.000000 (< 1)	0.00	

Save | Save and Close | Close

Edit Hazard

Use the "Name and Type" tab to edit the name and default units for the hazard. Use the remaining tabs to add model elements to the hazard.

A hazard requires at least one dose response model and one hazard metric to be included in a computed risk scenario.



Quick Links: Hazards

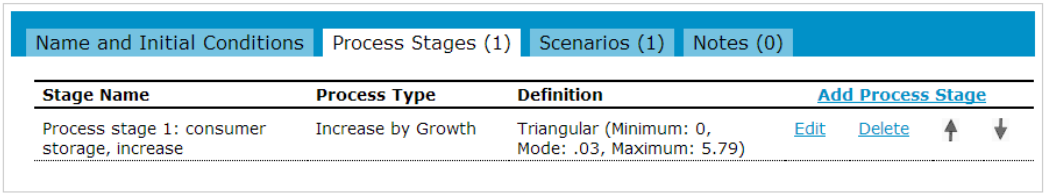
Edit Process Model

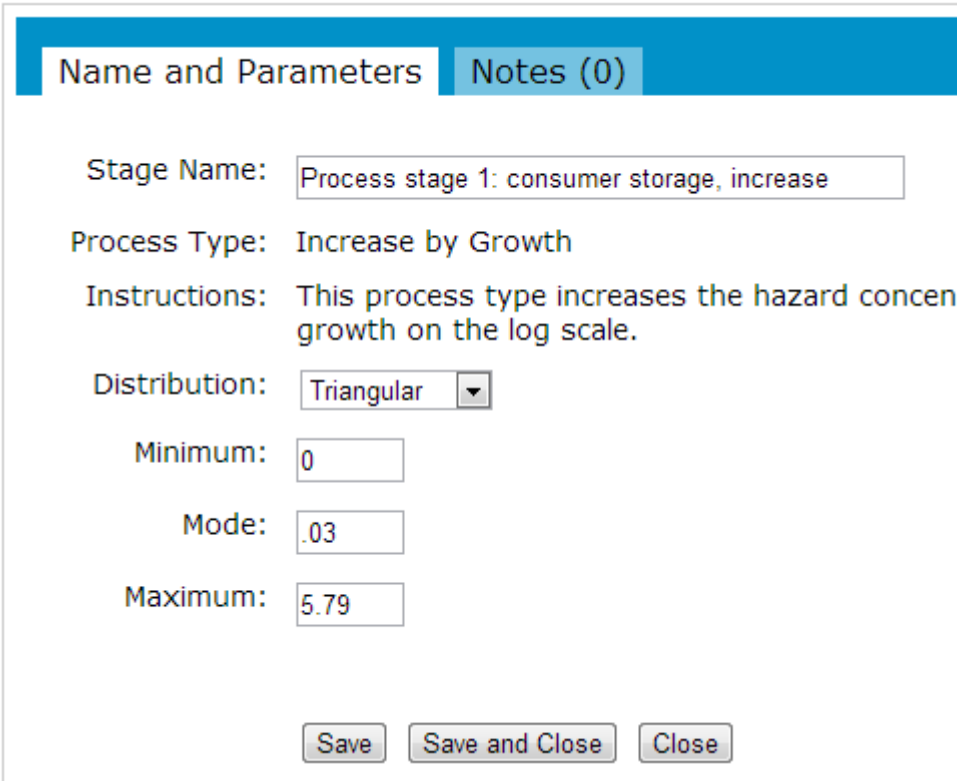
Use the "Name and Initial Conditions" tab to edit the model name and specify initial concentration, prevalence and unit size for the food-hazard combination.

Use the "Process Stages" tab to define the overall process model steps by adding/editing/deleting individual process stages in a defined sequence.

If the "Initial Units are Contaminated" checkbox is cleared, then any values for prevalence and concentration are ignored and the values are set to zero when the process model is evaluated.

If initial units are contaminated, the distribution provided for concentration must result in a mininum of one cfu/pfu per unit size specified.





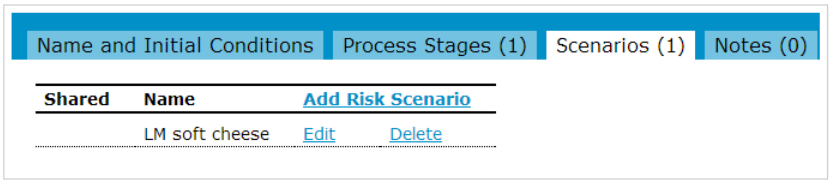
Edit Process Model

Use the "Name and Initial Conditions" tab to edit the model name and specify initial concentration, prevalence and unit size for the food-hazard combination.

Use the "Process Stages" tab to define the overall process model steps by adding/editing/deleting individual process stages in a defined sequence.

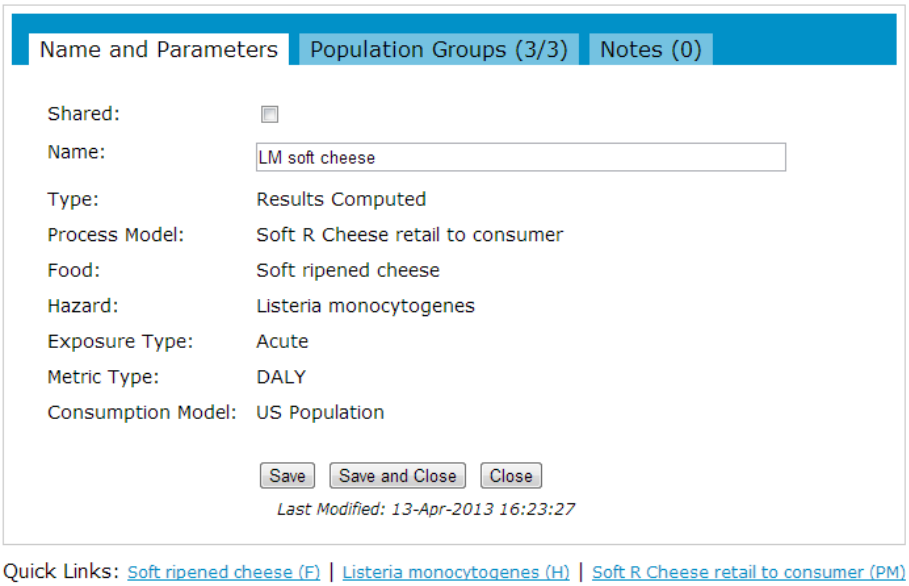
If the "Initial Units are Contaminated" checkbox is cleared, then any values for prevalence and concentration are ignored and the values are set to zero when the process model is evaluated.

If initial units are contaminated, the distribution provided for concentration must result in a mininum of one cfu/pfu per unit size specified.



Edit Risk Scenario

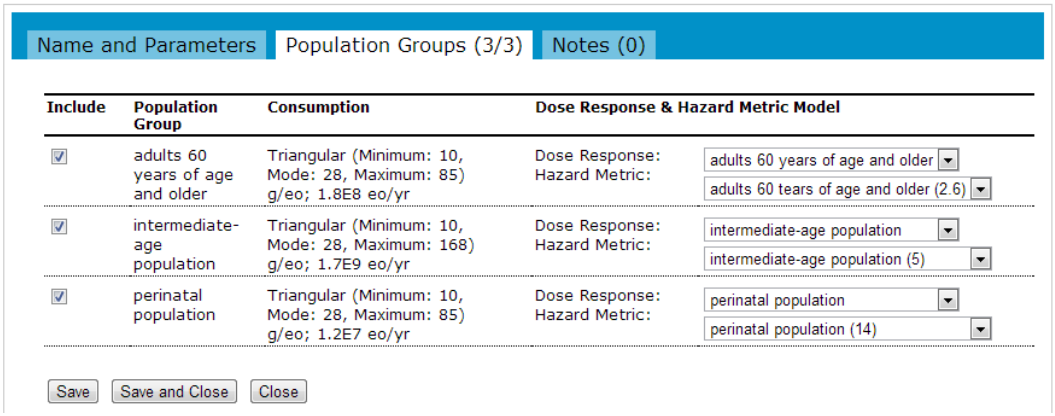
Use this page to change the name of the scenario, or to modify which population groups are include which dose response models and hazard metric values are used. At least one population group must included to run the scenario.



Quick Links: Soft ripened cheese (F) | Listeria monocytogenes (H) | Soft R Cheese retail to consumer (PM)

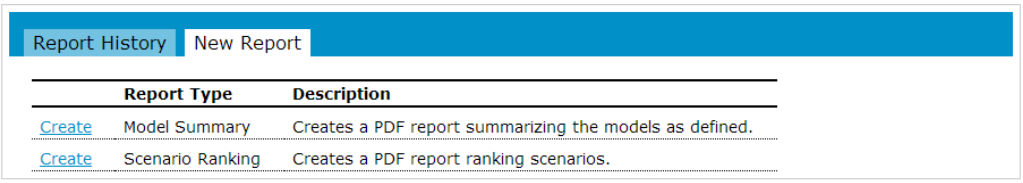
Edit Risk Scenario

Use this page to change the name of the scenario, or to modify which population groups are included and which dose response models and hazard metric values are used. At least one population group must be included to run the scenario.



Reports

Use the Report History tab to view pending reports still in the queue and to generate PDFs for completed reports. Use the New Report tab to create new reports.



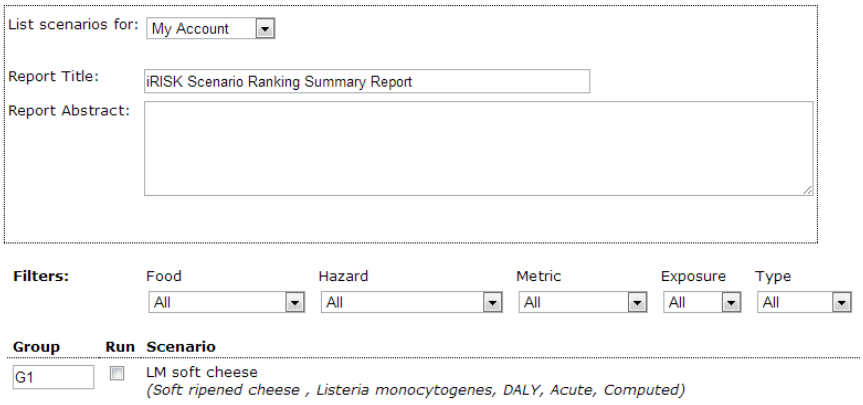
Rank Scenarios Report

The report title and abstract provided will be included in the report. Use the checkboxes to select which scenarios to include in the report and click "Generate Report For Checked". Or, click "Generate Report For All Listed" to select all.

If the list of scenarios is very long, use the filters to refine the list.

Any scenarios with the same Group text will be treated as a group during ranking (their individual results will be summed).

Clicking on the "Generate" buttons will submit the request to a queue. Use the Report History tab on the Reports page to view its status.



Group	Run	Scenario
G1	☑	LM soft cheese *(Soft ripened cheese , Listeria monocytogenes, DALY, Acute, Computed)*

1 scenarios submitted. Use the Report History tab to check the report status.

Generate Report For Checked | Generate Report For All Listed

Report History | New Report

Refresh

Completed Reports Filter by date: Last 2 Weeks Filter by status: All

Name & Abstract	Scenarios	Completed	Status	
iRISK Scenario Ranking Summary Report	1	13-Apr-2013 16:25:59	View PDF Details ☐ Notes ☐	Delete
iRISK Scenario Ranking Summary Report	1	13-Apr-2013 16:29:28	View PDF Details ☐ Notes ☐	Delete
iRISK Scenario Ranking Summary Report	1	13-Apr-2013 19:33:02	View PDF Details ☐ Notes ☐	Delete

iRISK Scenario Ranking Summary Report

Report Date: 2013-Apr-13 12:34:30

Disclaimer

FDA, JIFSAN and Risk Sciences International have taken all reasonable precautions in creating the iRISK quantitative risk assessment system (version 1.0) and the documentation accompanying it. The system is housed on the JIFSAN foodrisk.org website. However, FDA, JIFSAN and Risk Sciences International are not responsible for errors, omissions or deficiencies regarding the system and the accompanying documentation. The iRISK system and the accompanying documentation are being made available "as is" and without warranties of any kind, either expressed or implied, including, but not limited to, warranties of performance, merchantability, and fitness for a particular purpose. FDA, JIFSAN and Risk Sciences International are not making a commitment in any way to regularly update the system and the accompanying documentation.

Responsibility for the interpretation and use of the system and of the accompanying documentation lies solely with the user. In no event shall FDA, JIFSAN or Risk Sciences International be liable for direct, indirect, special, incidental, or consequential damages resulting from the use, misuse, or inability to use the system and the accompanying documentation.

Third parties' use of or acknowledgment of the system and its accompanying documentation, including through the suggested citation, does not in any way represent that FDA, JIFSAN or Risk Sciences International endorses such third parties or expresses any opinion with respect to their statements.

iRISK Scenario Ranking Summary Report

Hazard Metric: DALY

All reported summary values are per year. For chronic scenarios, results for the total lifecourse have been divided by the lifecourse duration (e.g. 70 years) specified for the population groups included in the scenario.

Group ID: G1 **Total DALYs per Year:** 19.2

Scenario	Lifecourse Duration	Eating Occasions or # of Consumers	Total Illnesses	Mean Risk of Illness	DALYs	Per Eating Occasion or Consumer
LM soft cheese (Acute, Computed)	N/A	1.89E+9	3.36	1.77E-9	19.2	1.02E-8

iRISK Scenario Ranking Summary Report

Summary for Risk Scenario: LM soft cheese

Group:	G1		
Hazard:	Listeria monocytogenes (Microbial Pathogen)	Scenario Type:	Results Computed
Food:	Soft ripened cheese	Exposure Type:	Acute
Process Model:	Soft R Cheese retail to consumer	Metric Type:	DALY
Consumption Model:	US Population		
Converged:	Yes (by 12000 samples)		

Process Model Details for: Soft R Cheese retail to consumer

Initial Prevalence:	0.0104	**Final Mean Prevalence:**	**0.0104**
Initial Concentration:	Triangular (Minimum: -1.39, Mode: -1.15, Maximum: 0.699) log10 cfu/g		
Initial Mean Concentration:	-0.338 log10 cfu/g	**Final Mean Concentration:**	**3.56 log10 cfu/g**
Initial Unit Size:	227 g	**Final Unit Size:**	**227 g**

			Post Stage Mean		
Process Stage	**Process Type**	**Definition**	**Concentration (log10 cfu/g)**	**Prevalence**	**Unit Size (g)**
Process stage 1: consumer storage, increase	Increase by Growth	Triangular (Minimum: 0, Mode: .03, Maximum: 5.79)	3.56	0.0104	227

Result Summary

Total Number of Illnesses: 3.36 **Total DALYs/Year:** 19.2

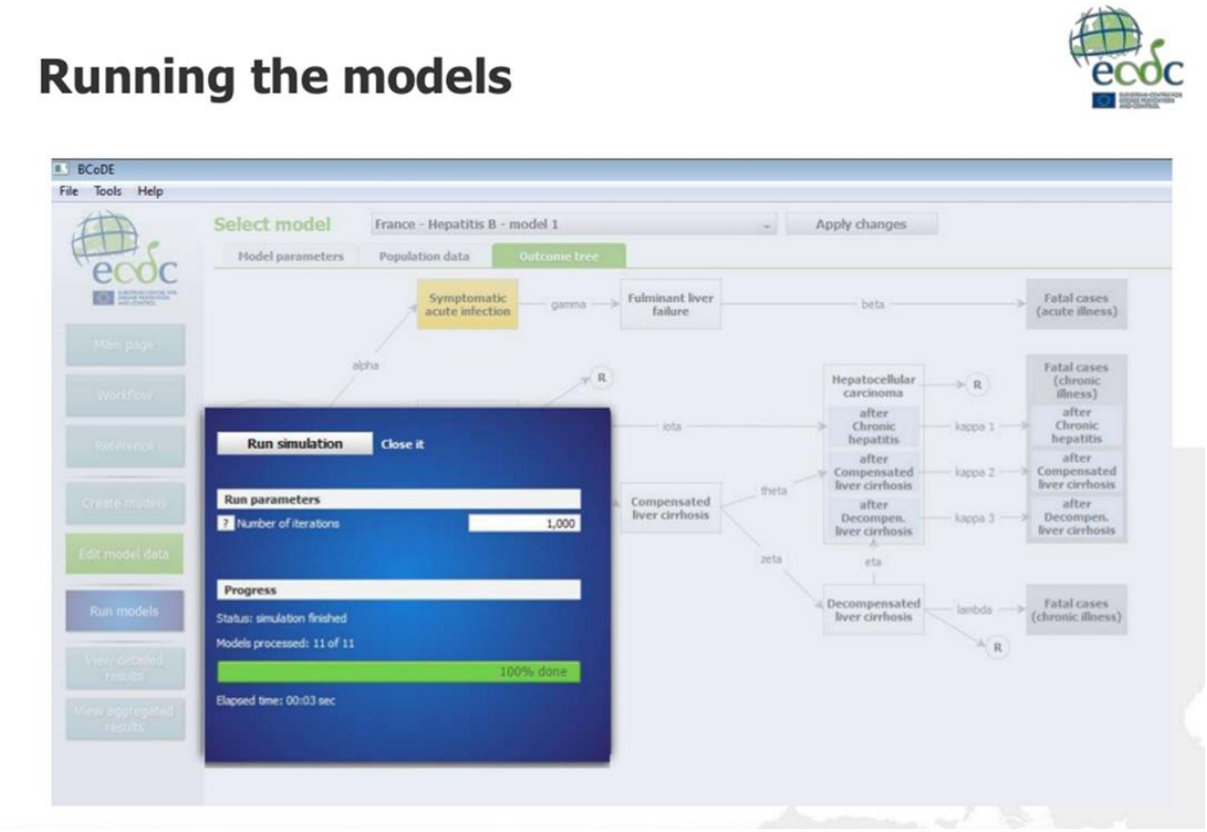
iRISK Scenario Ranking Summary Report

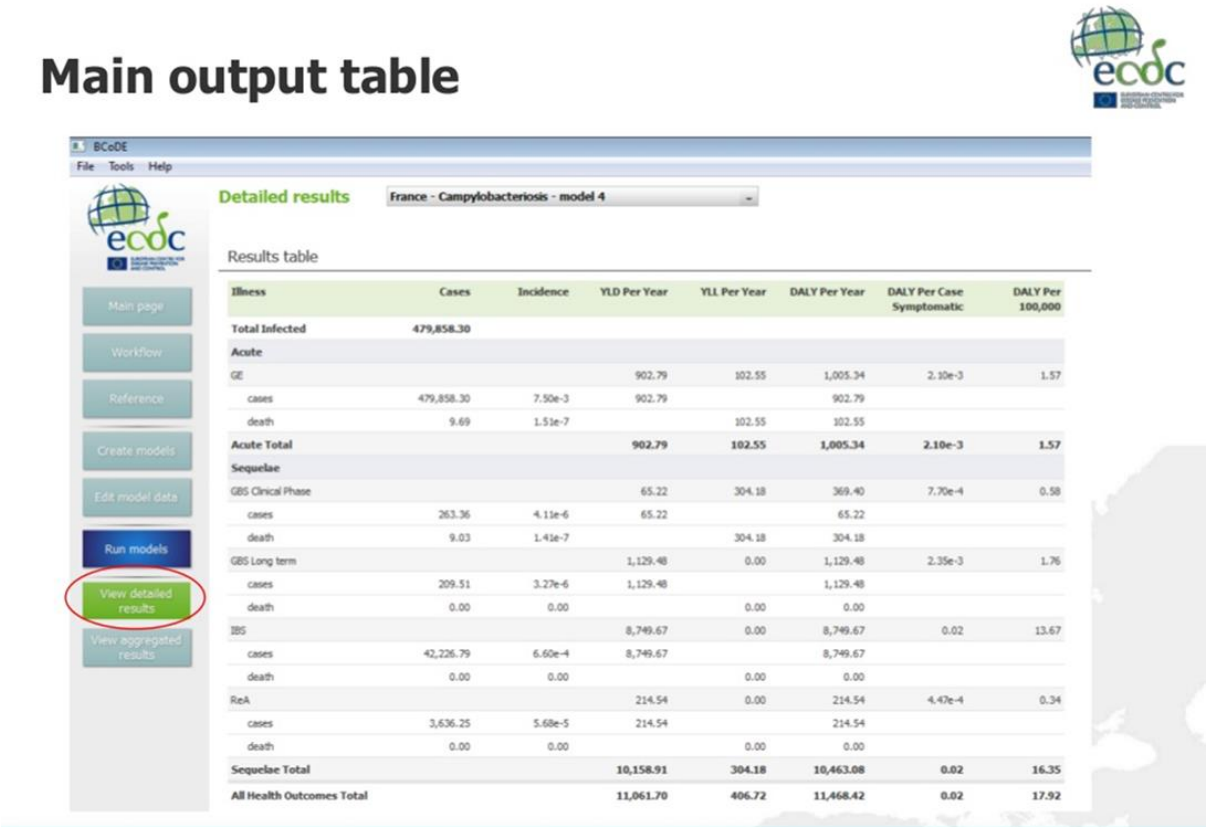
Detailed Results:

Population Groups	Consumption			Mean Probability of Illness	Number of Illnesses per year	Total Metric Per Year (DALYs)
adults 60 years of age and older	Triangular (Minimum: 10, Mode: 28, Maximum: 85) g/eo;1.8E8 eo/yr	Dose Response:	adults 60 years of age and older Dose unit: cfu (Exponential (r:8.39E-12);100%)	1.25E-8	2.25	5.84
		Hazard Metric:	adults 60 tears of age and older (2.6 DALYs)			
intermediate-age population	Triangular (Minimum: 10, Mode: 28, Maximum: 168) g/eo;1.7E9 eo/yr	Dose Response:	intermediate-age population Dose unit: cfu (Exponential (r:5.34E-14);100%)	1.41E-10	0.240	1.20
		Hazard Metric:	intermediate-age population (5 DALYs)			
perinatal population	Triangular (Minimum: 10, Mode: 28, Maximum: 85) g/eo;1.2E7 eo/yr	Dose Response:	perinatal population Dose unit: cfu (Exponential (r:4.51E-11);100%)	7.26E-8	0.871	12.2
		Hazard Metric:	perinatal population (14 DALYs)			

BCoDE

Examples of BCoDE results

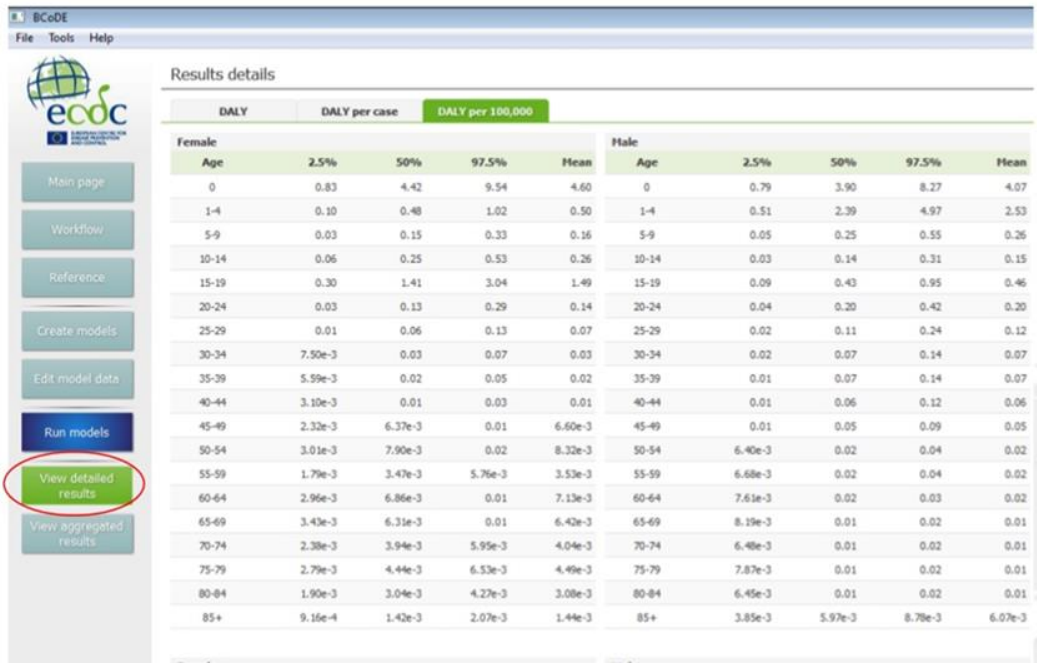




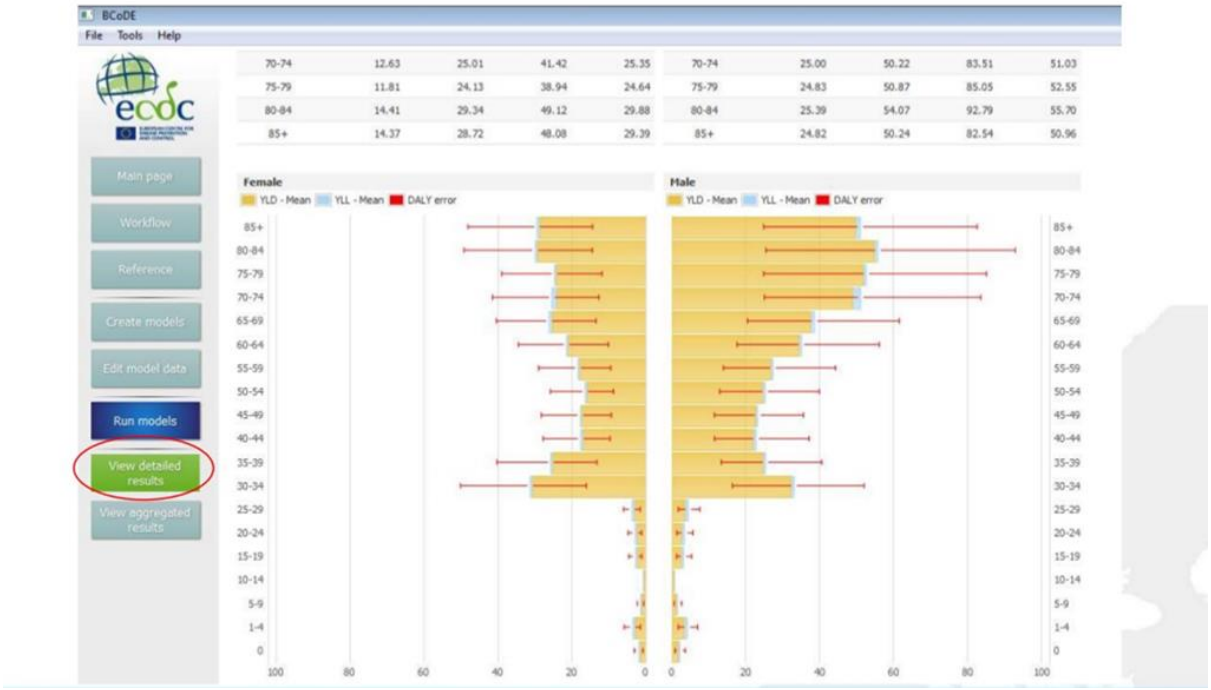
Illness	Cases	Incidence	YLD Per Year	YLL Per Year	DALY Per Year	DALY Per Case Symptomatic	DALY Per 100,000
Total Infected	**479,858.30**						
Acute							
GE			902.79	102.55	1,005.34	2.10e-3	1.57
cases	479,858.30	7.50e-3	902.79		902.79		
death	9.69	1.51e-7		102.55	102.55		
Acute Total			**902.79**	**102.55**	**1,005.34**	**2.10e-3**	**1.57**
Sequelae							
GBS Clinical Phase			65.22	304.18	369.40	7.70e-4	0.58
cases	263.36	4.11e-6	65.22		65.22		
death	9.03	1.41e-7		304.18	304.18		
GBS Long term			1,129.48	0.00	1,129.48	2.35e-3	1.76
cases	209.51	3.27e-6	1,129.48		1,129.48		
death	0.00	0.00		0.00	0.00		
IBS			8,749.67	0.00	8,749.67	0.02	13.67
cases	42,226.79	6.60e-4	8,749.67		8,749.67		
death	0.00	0.00		0.00	0.00		
ReA			214.54	0.00	214.54	4.47e-4	0.34
cases	3,636.25	5.68e-5	214.54		214.54		
death	0.00	0.00		0.00	0.00		
Sequelae Total			**10,158.91**	**304.18**	**10,463.08**	**0.02**	**16.35**
All Health Outcomes Total			**11,061.70**	**406.72**	**11,468.42**	**0.02**	**17.92**

Considering the impact of sequelae

More detailed information

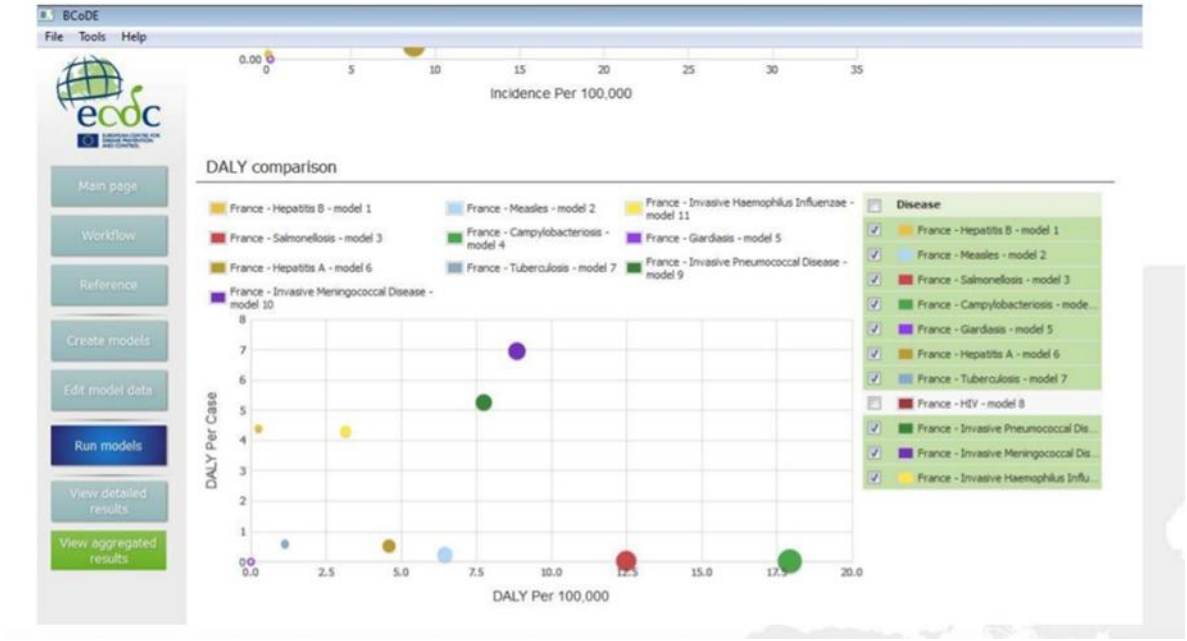


Demographic details with uncertainty levels

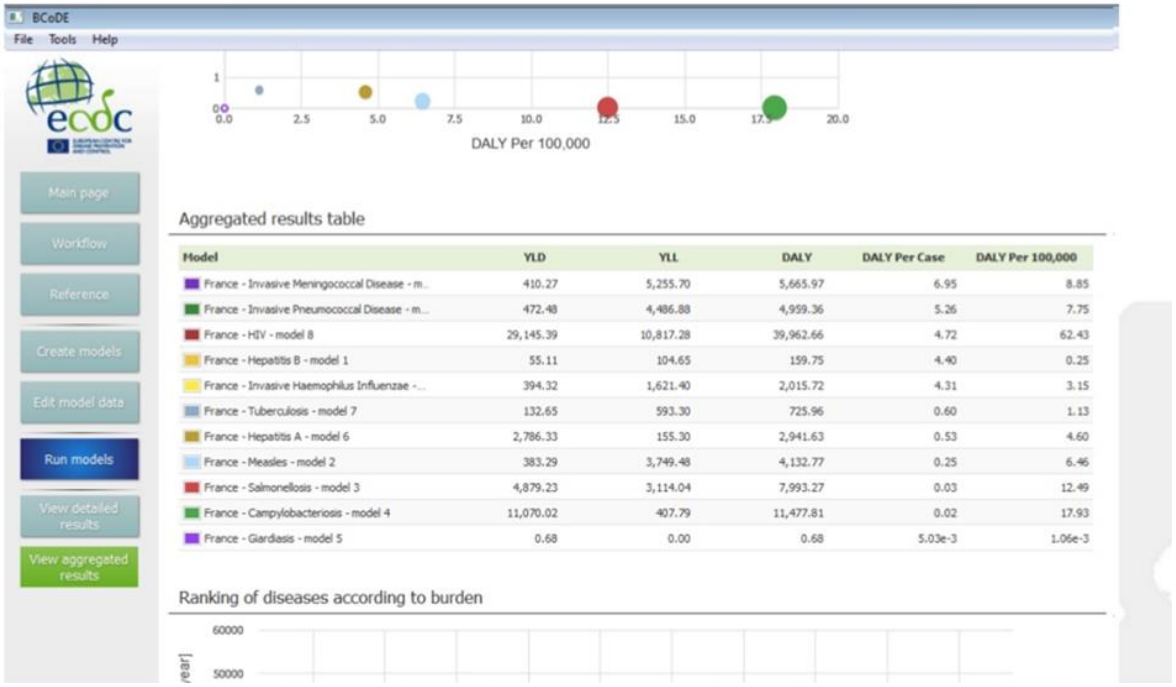


Tridimensional bubble charts

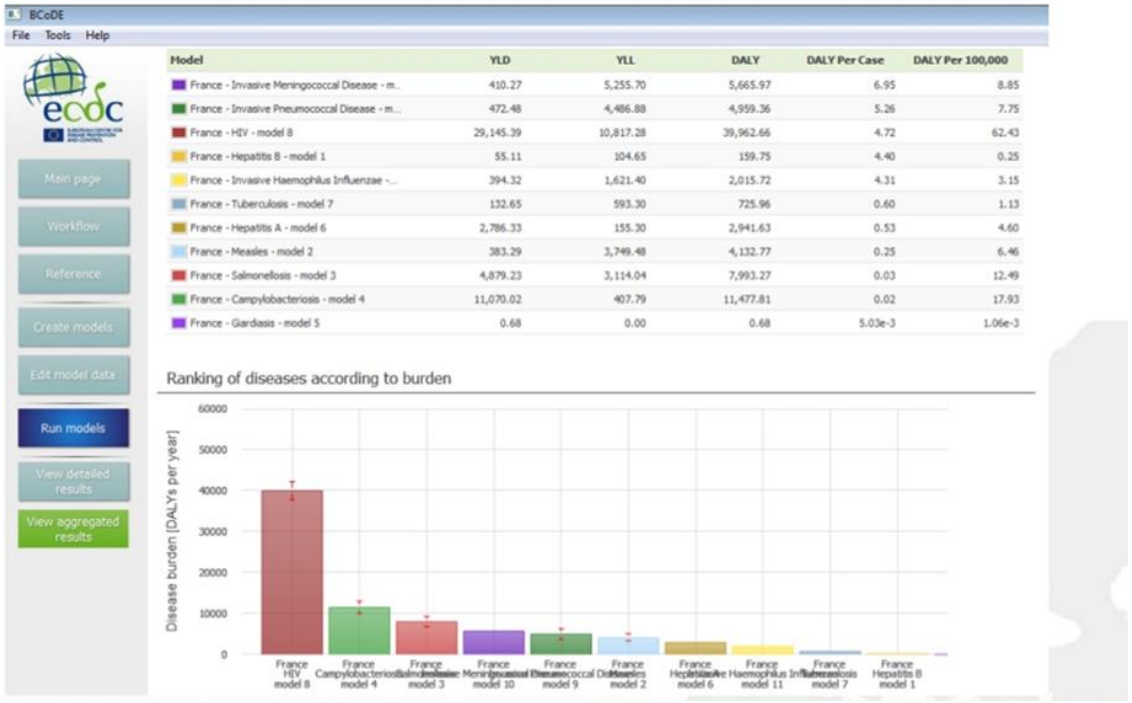
Tridimensional bubble charts



Ranking of diseases: the table



Ranking of diseases: the bar chart



ABBREVIATIONS

ADALY	Annual DALY
API	Annual probability of illness
APIll	Average probability of illness
BCoDE	ECDC Burden of Communicable Diseases in Europe
BIOHAZ Panel	EFSA Panel on Biological Hazards
CD	Communicable Disease
CDF	Cumulative distribution function
CFU	Colony-forming unit
DALY	Disability adjusted life years
ECDC	European Centre for Disease Prevention and Control
EFoNAO	EFSA food of non-animal origin
EFoNAO-RRT	EFSA food of non-animal origin risk ranking tool
EFSA	European Food Safety Authority
EU	European Union
FCID	Food Commodity Intake Database
FR	Frequency of consumption
GBD	Global Burden of Disease
$IllD_{50}$	The dose needed to cause illness in 50 % of exposed humans
JIFSAN	Joint Institute for Food Safety and Applied Nutrition
MS	Member State(s)
NHANES	National Health and Nutrition Examination Survey
NUSAP	numeral, unit, spread, assessment and pedigree
P^3ARRT	the pathogen–produce pair attribution risk ranking tool
PDF	Probability density function
PMP	Pathogen modelling program
QPR	Qualified Presumption of Risk
RI	Refrigeration index
RTE	Ready-to-eat
sQMRA	swift Quantitative Microbiological Risk Assessment
STEC	Shiga toxin-producing *Escherichia coli*
TESSy	The European Surveillance System
US	United States
US-FDA	United States Food and Drug Administration
WWEIA	What We Eat in America
YLD	Years lived with disability
YLL	Years of life lost due to premature mortality