EXTERNAL SCIENTIFIC REPORT

CRITICAL REVIEW OF METHODOLOGY AND APPLICATION OF RISK RANKING FOR PRIORITISATION OF FOOD AND FEED RELATED ISSUES, ON THE BASIS OF THE SIZE OF ANTICIPATED HEALTH IMPACT[1]

H.J. van der Fels-Klerx[1], E.D. van Asselt[1], M. Raley[2], M. Poulsen[3], H. Korsgaard[3], L. Bredsdorff[3], M. Nauta[3], V. Flari[4], M. d'Agostino[4], D. Coles[2], L. Frewer[2]

[1]RIKILT, Wageningen University and Research Centre, Akkermaalsbos 2, NL-6708 WB, Wageningen, the Netherlands; [2]University of Newcastle, School of Agriculture, Food and Rural Development, Agriculture Building, Newcastle Upon Tyne NE1 7RU, UK; [3]Technology University of Denmark, National Food Institute, Morkhoj Bygade 19, Soborg, Denmark; [4]Food and Environmental Research Agency, Sand Hutton, York, North Yorkshire, YO41, 1 LZ, UK.

ABSTRACT

This study aimed to critically review methodologies for ranking of risks related to feed/food safety and nutritional hazards, on the basis of their anticipated human health impact. An extensive systematic literature review was performed to identify and characterize the available methodologies for risk ranking in the fields of feed and food safety and nutritional hazards, as well as the socio-economic field. Risk ranking methods from the environmental field were studied as well to determine whether approaches used in this field could also be applied for ranking human health risks related to feed and food safety and nutritional hazards. The review used a predefined search protocol. It covered the bibliographic databases Scopus, CAB Abstracts, Web of Sciences, and PubMed over the period 1993-2013. All references obtained were stored into an Endnote database and evaluated for their relevance. All references deemed to be relevant were studied in–depth so as to characterize the risk ranking method described. Characteristics of each method were stored in an Excel database. The methods for risk ranking were then grouped into method categories, which were described in general. These groups included: risk assessment, comparative risk assessment, risk ratio method, scoring method, cost of illness, DALY/QALY, willingness to pay, multi criteria decision analysis, risk matrix, flow charts/decision trees and expert judgment methods. Based on the characteristics of the individual methods and the method categories, an overarching framework was developed for selection of the appropriate method(s) that could be used for risk ranking of feed and food related hazards, on the basis of human health impact. This framework has the format of a decision tool, with which – given the characteristics of the risk ranking question at hand - the most appropriate method(s) can be selected. Application of this overall framework to several case studies showed it can be a useful tool for risk managers/assessors to select the most suitable method for risk ranking of feed/food and diet related hazards, on the basis of expected human health impact.

DISCLAIMER

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

KEY WORDS

Risk prioritization, risk ranking, methodology, food safety hazards, contaminants, pathogens, nutritional hazards, health impact.

SUMMARY

Governmental and regulatory organisations consider the application of risk ranking as a possible basis for the prioritisation of the allocation of resources to mitigate feed and food related hazards and their anticipated public health impacts. This study aimed to give an overview of available risk ranking methods for this purpose. Each of the methods was critically reviewed to extract the potentials and limitations for the ranking of human health risks related to feed or food. The study covered toxicological, biological and nutritional health risks of well-known chemical substances, biological agents and nutritional components in food and feed. An extensive literature search was performed to identify the available methodologies for risk ranking in the fields of feed and food safety and nutritional hazards. Furthermore, literature on risk ranking approaches used in social sciences was included. Risk ranking methods from the environmental field were studied as well to determine whether approaches used in this field could also be applied for ranking human health risks related to feed and food safety and nutritional hazards.

The review performed was based on the principles of a systematic review, as much as possible. Beforehand, the literature search protocol was defined, including both search strings and criteria for evaluation of the literature references. Then, the literature review was performed by applying the predefined search protocol, using four bibliographic databases: Scopus, CAB Abstracts, Web of Sciences, and PubMed over the period 1993-2013. All references obtained were stored in an Endnote database. Then, the references were evaluated using the evaluation criteria. Based on title, abstract and keywords, references were categorized as relevant, maybe relevant or not relevant. Full texts were retrieved for the references in the maybe relevant group and based on the evaluation criteria these were then grouped into the relevant or not relevant group. Each relevant reference was studied in–depth to obtain a detailed evaluation of the characteristics of the particular risk ranking method described. This critical assessment covered topics like the type of tool, field, scope, application area (type of hazards), metrics (type of method), model structure, data requirement, data collection and data integration. Results of the critical evaluation were stored in an Excel database thus containing detailed information on all selected studies in which a risk ranking method was applied.

The various methods for risk ranking were grouped into method categories, including: risk assessment, comparative risk assessment, risk ratio method, scoring method, cost of illness, DALY/QALY, willingness to pay, multi criteria decision analysis, risk matrix, flow charts/decision trees and expert judgment methods. These categories of methods were studied for their general characteristics of scope, application area, approach, and strengths and weaknesses, and perspective for use by stakeholders.

Based on the results from the evaluation of the relevant references and the general evaluation of the categories of risk ranking methods, an overarching framework has been developed for selection of the appropriate method(s) that could be used for risk ranking of feed and food related hazards, considering the impact on human health. This framework has the format of a decision tool, with which – given the characteristics of the risk ranking question at hand - the most appropriate method(s) can be selected. The feasibility of the use of the overall framework for selection of appropriate methods was then evaluated by application of the framework to some pre-selected case studies. In consultation with EFSA, case studies were defined beforehand for several microbiological, chemical and nutritional hazards. Results from the feasibility study showed that, usually, more than one method can be applied for the risk ranking task at hand. Each of the methods has its own strengths and weaknesses. Further analysis of available data is then needed as well as insight into the risk manager's preferences in order to select the preferred method. The decision tool is expected to be valuable for EFSA and other risk managers/assessors as it helps to select the most appropriate methods for risk ranking of feed/food and diet related hazards, on the basis of human health impact.

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY EFSA

Risk ranking is regarded as a possible way for governmental and regulatory organisations to prioritise the allocation of resources to mitigate the effects of hazards and anticipated public health impacts. Previous studies, e.g. Van Kreijl et al. (2004), illustrated that the implementation of risk ranking approaches for prioritization of food and feed issues could be feasible and beneficial. Considerations should also include the legal obligations and deadlines, urgency (as indicated by the originator), level of public interest, economic importance of the subject or issue, and prioritization of the workload. With an optimal use of all information, it would be possible to assist EFSA in ranking risks to prioritise the identified issues, on the basis of the size of the anticipated public health impact. A suitable risk ranking methodology could form the starting point for conducting a subsequent agency-wide collaborative study (which would go beyond the scope of this call) to build the food related health risks database for Europe.

Before EFSA would be able to implement a ranking approach, it needs to build an evidence base, which would ideally examine toxicological, biological and nutritional health risks, where possible, using a common denominator (such as, but not limited to, Disability Adjusted Life Years, DALY, or Quality Adjusted Life Years, QALY). Such an evidence base would also need to address uncertainties and utilise social science risk analysis approaches and tools to ensure a comprehensive coverage of public concerns.

TERMS OF REFERENCE AS PROVIDED BY EFSA

The scope of this work is to **critically review** the potential and limitations of methodologies, tools and proposals for ranking of risks associated with diet and health, food and feed, including social science risk analysis approaches, related to:

- the toxicological, biological and nutritional health risks of well-known chemical substances, biological agents and nutritional components in food and feed;

- examples where these may be assessed based on existing and/or new common denominators.

This contract/grant was awarded by EFSA to: RIKILT, as part of Stichting DLO, Wageningen, the Netherlands

Contractor/Beneficiary:

Contract/grant title: "Critical review of methodology and application of risk ranking for prioritisation of food and feed related issues, on the basis of the size of anticipated health impact"

Contract/grant number: OC/EFSA/SCOM/2013/01.

1. Introduction and Objectives

Risk ranking of feed/food safety and nutritional related health risks is generally recognised as the starting point for risk-based priority setting and resource allocation, as it permits policymakers to allocate their resources on the most significant public health problems. It is frequently defined as the analysis and prioritization of the combined probability of food contamination, consumer exposure and the size of anticipated public health impact of certain foodborne hazards. It is, thus, the combination of the probability that a hazard may occur in a food product and the effect of the hazard on human health. Risk ranking has been applied in establishing food safety monitoring programs and has shown to increase the efficiency of monitoring and to decrease inspection costs (Baptista et al., 2012; Presi et al., 2008; Reist et al., 2012). Risk ranking may thus help to efficiently prioritize hazard-food combinations to be studied.

To date, various risk ranking methods are available that prioritise risks through attribution of illness, disease burden and costs (Van Asselt et al., 2012). Methods vary from qualitative to quantitative methods (Cope et al., 2010; Van Asselt et al., 2012). An example of a quantitative method is a microbiological risk assessment, which includes a hazard identification step, an exposure assessment, hazard characterisation and risk characterisation. Such methods provide detailed information on the risks involved and usually include uncertainty analysis. One of these quantitative methods is iRISK, which is a comparative risk assessment system for evaluating and ranking food-hazard pairs, developed by FDA (see http://www.foodrisk.org). However, quantitative methods are also very elaborate. As a results, several semi-quantitative methods have been developed such as Risk Ranger (Food Safety Centre, 2010; Ross and Sumner, 2002), the swift QMRA tool (Evers and Chardon, 2010) and P^3ARRT (Anderson et al., 2011). These methods determine the relative risk based on the product and hazard involved as well as the processing steps. Both qualitative and (semi-)quantitative methods may involve expert elicitation to determine severity and probability of occurrence of nutritional or feed/food safety hazards, or the use of focus groups to determine societal concerns and levels of interest from various groups of individuals.

Risk ranking may also be based on the use of risk matrices combining probability on the x-axis with severity on the y-axis (FAO, 1998). This approach can be further elaborated using risk assessment nomograms, such as described by Barendsz et al. (1997). Probability, in this case, is divided into likelihood of occurrence and manufacturing control, which is then multiplied with the severity of the effect to come to a qualitative indication of the risk. The various items are scored on a scale from high to low. The most comprehensive nomogram is developed by the UK Food Standards Agency (Lenartowicz and Michie, 2002a, 2002b). On top of the previously described items, the approach incorporates a consumer part including consumption patterns and population affected.

As part of the ranking methods, metrics in the area of Health Adjusted Life Years (HALY) typically involve calculating the combined incidence and weights (potentially developed from an individual's experience of suffering from a condition) for different health states or conditions with estimates of life expectancy. This results in a measure of the health-related quality of life (HRQoL) and an indicator for the period over which that quality of life is experienced. Prominent examples in risk ranking include

QALYs (Quality Adjusted Life Years) and DALYs (Disability Adjusted Life Years). DALY is estimated as the sum of years of potential life lost due to premature mortality and the years of productive life lost due to disability (Polinder et al., 2012). QALY indicates life years gained due to medical interventions. While QALYs aggregate the actual health quality over time, DALYs aggregate the loss of health compared to perfect health (Bogaardt et al., 2004).The QALY loss associated with an adverse health state is measured as the difference between QALYs with and without the condition. Different methods have been applied to estimation of utility values of health states across different domains linked to quality of life experienced (such as pain, anxiety and perceived quality of life) which are weighted using different methodologies (such as the Visual Analog Scale, Person Trade-Off, Time Trade-Off, Standard Gamble and Pairwise Comparison). The DALY/QALY concept may incorporate medical and economic consequences of food risks but they do not explicitly take societal impacts into account. Therefore, additional social elements may need to be included that can assess consumer perceptions and attitudes (Cope et al., 2010; Dreyer et al., 2010; Frewer et al., in press; van Dijk et al., 2011). Also, monetary risk metrics may be used in risk prioritization, such as willingness to pay or willingness to accept, which measure what individuals would be willing to pay to obtain health improvements or avoid adverse health states.

One of the broader risk ranking approaches, also sometimes considering socio-economic factors, is the Multi-Criteria Decision Analysis (MCDA) methodology. MCDA uses criteria such as epidemiology, prevention/control, effects on economy/trade, zoonotic characteristics, and societal effects. Expert elicitation is used to weigh the various criteria (Humblet et al., 2012) or provide the necessary information (i.e. strict preferences) from which weights of the criteria can be inferred (Neslo, 2011). A number of possible (modelling) approaches on how to construct, and possibly aggregate, individual preferences have been published (French et al., 2009; Lichtenstein and Slovic, 2006). A novel MCDA approach that employs discrete preference data and applies probabilistic inversion to quantify, with uncertainty, the weights in the model has been developed by Neslo (2011), and has been successfully applied to several fields, amongst them, nanotechnology enabled food products (Flari et al., 2011), environmental policy making (Teck et al., 2010), ranking preferences of regulatory actions to ensure safety of products and applications based on emerging sciences and technologies (Flari and Chaudhry, 2012), and ranking of emerging zoonoses (Havelaar et al., 2010).

Most of the risk ranking approaches available to date have been established for application to microbiological hazards. However, chemical hazards and nutritionally-related health risks may be ranked analogously with these methods (van Asselt et al., 2013). In order to determine which methods are most suitable for ranking food and feed related issues in general, it is important to follow a structured approach of identifying and evaluating the available methods, with transparency and objectivity being important factors (van Asselt et al., 2013).

The aim of the current study was to perform a review of current risk ranking methodologies for prioritisation of food and feed related issues, on the basis of size of anticipated health impact, to perform a systematic evaluation of the available methods, and to identify possible gaps in existing methods.

This overall aim was broken down into the following three specific objectives:

Objective 1: Extensive literature search and data management: to perform an extensive literature review focused on methods for risk ranking for application in diet and health, food and feed safety, including those methods developed in natural/life and socio-economic sciences, risk analysis and food safety governance, and to handle the obtained data in an non-subjective and transparent way.

Objective 2: Information clustering: to evaluate the identified methods/tools for their characteristics, weaknesses and strengths, data resources, fields of applications, and to group, as much as possible, the various methods/tools, based on these aspects and their functioning in several case studies.

Objective 3: Applicability for EFSA: to critically compare all methods/tools for risk ranking for prioritisation of food related issues, to design a conceptual framework for an overarching risk ranking approach, and to develop recommendations for future tool development, and use by EFSA.

2.   Materials and Methods

2.1.   Protocol for literature search

A structured literature review was conducted aiming to identify the available risk ranking methodologies used in the fields of feed and food safety and nutritional hazards, on the basis of the size of anticipated health impact. It included those methods developed from both natural/life and socio-economic sciences and food safety governance. The literature review followed, as much as possible, the principles of a systematic literature review, as described by EFSA (2010). A protocol for the structured literature review was defined beforehand. It included both **criteria for searching** relevant literature (search strings) and **criteria for evaluation** of obtained literature references (applying the search criteria). First, preliminarily versions of the protocol and the criteria for evaluation were defined. These were then used for test runs and results were discussed with EFSA. Based on the first results and the discussion, both the criteria for searching and for evaluation were adapted. As part of the protocol, a template of extraction of literature references and their evaluations was made.

2.1.1.   Draft literature search protocol

a)   Criteria for searching literature:

The search strategy consisted of three major steps, each designed to search titles and subject headings. Combinations of search strings were used, starting with a broad screening on methodology for risk ranking and prioritisation methods and tools in the field of feed and food related issues (step 1), then narrowing down the methods relating to size of anticipated impact on human health (step 2), and finally focusing specifically on chemical hazards, biological hazards and nutritional components as well as social issues related to feed and food (step 3). The strategy steps and proposed search strings are as follows:

Step 1 –captured titles/subject headings that studied methods and tools for risk ranking and prioritization related to feed and food issues. This step included the following search terms ("*" indicates that the search term includes all words containing additional characters following the "*" symbol).

(RISK* or HAZARD*) and (RANK* or METHOD* or NOMOGRAM* or MATRIC* or DECISION TREE* or PRIORI* or ANALYS* or MC*A or MULTI-CRITERI*) and (FEED or FOOD* or AGRI* or AGRO*)

AND

Step 2 –captured titles/subject headings that investigated risk ranking and prioritisation methods on the basis of anticipated health impact. This step included the following search terms:

[for FERA, RIKILT, DTU]

DALY* or QALY* or HALY* or DISEASE BURDEN* or HUMAN HEALTH* or TOX* or ILLNESS* or COST* or SEVER* or ADI* or TDI*

[for UNEW]

("socio* impact" OR "econ* impact" OR WTP OR cost*)

AND

Step 3 –captured titles/subject headings that investigated specific application fields of biological agents, chemical hazards and nutritional components in feed and food, as well as social science issues related to food hazards, from consumer and governance perspectives. This step included the following search terms:

1. Microbiological hazards:
   a. ZOONOS* or
   b. MICROBIOL* or
   c. GENE* or
   d. PATHOGEN* or
   e. QMRA

2. Chemical/toxicological hazards:
   a. NANO* or
   b. CHEMIC* or
   c. ANTIBIOTIC* or
   d. DIOXIN* or
   e. HEAVY METHAL* or
   f. CARCINOGEN* or
   g. PESTICIDE* or
   h. HORMONE* or
   i. MYCOTOXIN* or

      j.   TOXIN* or

      k. BIOTOXIN*

3. Nutritional hazards
   a. NUTRI* or
   b. ANTI-NUTRI* or
   c. FLAVO* or

4. Societal impact:

   a. ATTITUTE* or
   b. PERCEP* or
   c. ACCEPT* or
   d. OPINION* or
   e. VIEW* or
   f. BEHAVIOUR* or
   g. ECON* or
   h. SOCIETAL* or
   i. SOCIA* or
   j. EMPLOY* or
   k. COMMUNICAT* or
   l. DIALOGUE* or
   m. ENGAGE* or
   n. PARTICIP* or
   o. GOVER* or
   p. LEGAL* or
   q. LAW* or
   r. REGUL*

2.1.2. Evaluation criteria

The references judged to be relevant for the study objectives were evaluated for eligibility and quality of the described research. The following preliminarily set of criteria were defined for evaluation of the literature results:

1. Relevance of the reference for the objective of the literature review;

2. International peer-reviewed journals and reports, and rank of the journal;

3. Methods were well described, (semi-)quantitative or qualitative, user-friendly, transparent, structured, and objective;

4. Methods were applicable in wider decision making schemes/frameworks;

5. In case of reports, they should originate from well-known, highly-respected governmental bodies or research organisations.

The preliminarily search strings and criteria for evaluation were discussed with the entire project team and with EFSA. Also, pilot searches utilising the protocol were conducted and - based on the test results - the protocol was further refined. This iterative process ultimately led to the final version of the literature search protocol.

2.2. Literature review

2.2.1. Review methodology

A literature search was conducted with the aim of identifying methods for risk ranking in the fields of feed and food safety and nutritional components, on the basis of the size of anticipated health impact, including those methods developed in natural/life and socio-economic sciences and food safety governance. The review followed the methodology described given below, applying the final literature search protocol.

Methodology for the literature search:

a. The review focused on published literature from the latest 20 years (1993-2013). It used the following four bibliographic databases: Web of Science, Scopus, PubMed, and CAB Abstracts. In addition to these bibliographic databases, the general search engine Google was used, in particular to search study reports ('grey literature') from relevant international and national organisations, authorities, and agencies (e.g., EFSA, EMA, WHO/FAO, FDA, Health Canada, OECD). Including only English language references can introduce language bias. Large bibliographic databases often include a smaller number of non-English language journals. Therefore, any references in the German, French, Italian, Spanish or Dutch languages (but with English search strings) were also included in the review, if deemed relevant, such as to minimize potential language bias.

b. The final set of search strings was applied, leading to an initial set of search results. Duplicates within this set, as a consequence of using four different bibliographic databases, were removed.

c. The references of the initial set of search results were screened for their relevance to the project objectives, using the final set of evaluation criteria. A two-tier approach was used. In tier 1, the applicability of each reference to the review objective was determined by examining the **title, abstracts and key-words** of each reference. Based on this evaluation, the references were sorted into one of three categories:

- *Relevant for this study*: Based on title/headings, abstract and key-words the reference was included.

- *Maybe relevant for this study*: Based on title, abstract and key-words, it remained unclear if the reference was relevant for the study. Later on, the full text of the reference was obtained to

assist in this determination. References ultimately determined to be relevant were included whereas references determined not to be relevant were moved to the third category.

- *Not relevant for this project*: Based on title, abstract and key-words the reference was determined to be out of the scope of this study.

All references evaluated were categorized in each of the three groups indicated in the Endnote file. An inter-observer reliability check was done with a subset (10%) of both selected and excluded references.

d. In tier 2, the underlying references/reports for references in the relevant and maybe relevant group of the Endnote database were retrieved. By reading the **full texts**, there references/reports were evaluated for their relevance to the field of interest and their quality using the evaluation criteria. When deemed relevant, the reference was kept in the group *Relevant* in the Endnote database. When deemed not relevant, the reference was removed to the group *Not relevant* in the Endnote database.

e. Citations used in the reports/references of the final Endnote database were screened for additional relevant references, published after 1993, applying step c and d to these sources (snowballing citation).

f. Furthermore, the network of the project team was used for identification of (maybe) relevant references not yet included. These references were then searched for, and step c and d were applied to these sources, published after 1993.

The evaluation of the methods presented in the relevant literature references was performed by different individuals of the project team. Certain (chosen at random) literature references were used as *doubles* in order to gain a view of the variability of the evaluation results of the different individuals.

2.3.   Database for literature results

All references, either initially or ultimately determined to be relevant and eligible for the study objectives (see section 2.2), were entered into a final Endnote database. This database thus contained a highly relevant and quality controlled list of literature references, of direct relevance to the research objectives. All references that were evaluated but deemed not to be relevant were stored into a separate Endnote database.

2.4.   Database for risk ranking approaches

Using the references in the final Endnote database, an in-depth evaluation of the characteristics of the risk ranking method(s) described in each reference was performed. This critical assessment covered topics such as the type of tool, application area (type of hazards), metrics (type of method), model structure, model variables, data collection and data integration.

A format for an excel sheet for storing the information obtained from the relevant references was defined before-hand. The template proposed by the BIOHAZ panel was used as the starting point for defining the format (EFSA, 2012c); this template was modified to the specific objectives of the current study. The defined format contains columns for information about the reference, such as author names, title, abstract and journal, volume and page numbers, as well as columns for storing results from the critical assessment of the risk ranking methods. Different columns were created for:

- The type of tool,

- Field of application: microbiological, chemical, nutritional hazards,

- Scope/What was ranked, e.g. specific food products,

- Specific application area, e.g., pesticides,

- Metrics, i.e., the type of method, with different sub-columns for each method category,

- Model structure: quantitative, semi-quantitative or qualitative,

- Data requirement, describing the model variables, such as human population data, ADI, MRL, microbial numbers etc.

- Data collection, describing how the necessary data were collected, which data sources were used,

- Data integration, describing how data was integrated in the application described in the reference.

Each relevant reference from the Endnote file was then retrieved, and its full text was studied to characterise the particular risk ranking method described at hand. Moreover, the various risk ranking methods were grouped into different functional categories by methodology. Results of the evaluation of the methods were stored in the excel database. Unique rows were used for each of the relevant references. In this way, the results from Endnote can be linked to the results stored in the evaluation sheet.

2.5.   Overall conceptual framework

The references describing studies referring to the same category of risk ranking method were considered together. They were used to describe the particular category for risk ranking methods in a more general way using common, qualitative characteristics. Method categories were evaluated and described for: scope, application area, approach, and strengths and weaknesses, and perspective for

use by stakeholders. In this more general evaluation, reviews on risk ranking methods and other relevant literature were consulted as well.

Using the results from the evaluation of the relevant references, stored in the excel database, and the general evaluation of the categories of risk ranking methods, a framework was developed for selection of the appropriate method(s) that could be used for risk ranking of feed and food related hazards, considering the impact on human health.

2.6.   Feasibility study

The feasibility of using the overall framework for selection of appropriate methods in response to various generic risk ranking problems was evaluated by application of the framework to some pre-selected case studies. Case studies were defined for microbiological, chemical and nutritional hazards in consultation with EFSA. This resulted into the following cases:

Chemical/toxicological:

- *Fusarium* mycotoxins,

- pyrrolizidine alkaloids,

- botanicals in dietary supplements.

Microbiological:

- viruses in food of non-animal origin,

- pathogens in food of animal origin.

Nutritional:

- mortality effects of 12 modifiable dietary, lifestyle and metabolic risk factors

Chemical, microbiological and nutritional hazards:

- Comparing hazards from different fields using a common denominator

For each case, the aim of the risk ranking exercise was clearly specified: was it meant to gain scientific insight, to determine scientific budgets or as input to monitoring programs etc. In order to determine the preferred risk ranking method, the risk assessor should communicate with the risk manager to determine his/her preferences, prerequisites and assumptions. These should be clarified before the overall framework can be used. Therefore, for each of the cases presented above, a case description was defined a priori, in which the requirements and assumptions were given.

Then, the overall framework was used to select the appropriate method or methods to be used for risk ranking for the case at hand, given the case characteristics and requirements of the end-user, matching with the characteristics of the methods. For one microbiological case and one chemical case study, the actual risk prioritization, defined in the case description, was performed. To this end, the necessary data were collected from literature.

3. Results

3.1. Literature search protocol

Results from pilot searches using the preliminarily set of search strings indicated that the search strings needed to be adjusted. On the one hand, some specific search strings were added, on the other hand, the search needed to be refined. The final search strings used are listed below.

Step 1:

TOPIC = (risk* OR hazard*) AND
TITLE = (categor* OR rank* OR method* OR nomogram* OR matric* OR decision* OR priori* OR analys* OR mc*a OR multi-criteri* OR assessment*) AND
TOPIC = (feed OR food* OR agri* or agro*OR environ*) AND

Step 2:

[for FERA, RIKILT, DTU]
TOPIC = (disease* OR human health* OR *tox* OR illness* OR cost* OR sever* OR adi* OR tidI* OR epidemiol* OR BoD OR wtp OR incidence OR prevalence)

[for UNEW]
TOPIC = ("socio* impact" OR "econ* impact" OR WTP OR cost* OR WTA)

Step 3:

[for FERA]

Microbiological hazards:

TITLE = (zoonos* OR microb* OR gen* OR pathogen* OR qmra OR "antimicrobial resistance" OR parasite* OR virus* OR bacteria* OR micro*rgan* OR prion* OR TSE* OR QRA) AND

NOT = benefit*

[for RIKILT]

Chemical/toxicological hazards:

TITLE = (nano* OR chemic* OR antibiotic* OR dioxin* OR "heavy metal*" OR carc* OR pesticid* OR "plant protection product*" OR hormon* OR mycotoxin* OR phytotoxin* or phycotoxin* or marine biotoxin* OR Biocid* OR *contam* OR *pollutant* OR Melamin* OR Acrylamid* OR PCB* OR Residu* OR Endocr* OR Mutag* OR Botanic* GMO* OR "Genetic* modif*" OR "Novel protein*" OR Allerg* OR Insecticid* OR Acaricid* OR Herbicid* OR Fungicid* OR "plant growth regulat*" OR POP OR POPs OR Persistent* OR *accumul*) AND

NOT = benefit*

[for DTU]

Nutritional hazards

TITLE = (*nutri* OR *diet* OR bioavail* OR *supplement* OR "Novel protein*" OR Fortification* OR "Novel food*" OR Allerg*) AND

NOT (toxic* OR microbial* OR chemic* OR socio* OR benefit*)

DALY/QALY concept:
TOPIC = (daly* OR qaly* OR haly* OR HRQL* OR HALE) AND
NOT = benefit*

[for UNEW]

Societal impact:

TOPIC = ("focus group*" OR survey* OR interview* OR public* OR "expert analys*" OR *attitud* OR *percep* OR Willingness* OR *Soci* OR Determ* OR Cultur* OR Tradition* OR Typic* OR Consumer* OR Ethic* OR accept* or opinion* or view* or behaviour* or behavior* or employ* or communicat* or dialog* or engage* or particip* or gover* or legal* or law* or regul*) AND

NOT: religious* or halal* OR benefit*

Application of these search strings led to the following numbers of initially retrieved references (after removing most of the duplicates):

- 6322 references were retrieved via the search strings for chemical/toxicological hazards,

- 3092 references were retrieved via the search strings for microbiological hazards,

- 1103 references were retrieved via the search strings for nutritional hazards,

- 115 references were retrieved via the search strings for DALY/QALY concept,

- 3578 references were retrieved via the search strings for socio-economic methodology.

For four specific ("examples") topics - being hormon*, fortification*, parasite*, consumer* - the evaluation criteria were applied to the retrieved set of references, in order to evaluate at an early stage if the selected evaluation criteria worked well. Based on this exercise, the evaluation criteria were not changed. However, the first evaluation criterion ("relevance of reference of the objective of the review") was augmented by adding more specific criteria for judging the relevance of references to the current study objectives. These included:

a) References discussing prioritisation/ranking methods for human health risks and/or,

b) References describing risk prioritization/ranking methods applied for environmental/ecological risks and/or,

c) References to risk prioritization, risk analysis, risk assessment methods and/or risk modelling included in abstract and/or,

d) Any relevance of the work for application to human health, including references on drinking water and/or,

e) Abstract indicates socio-economic research methodology is employed.

Criteria for excluding references were:

- References discussing only parts of a method, such as references dealing with presence of chemical hazards, analytical methods, and/or references about toxicity studies. These are all parts of a risk assessment and/or,

- References addressing non-human related aquaculture and non-human related animal health.

As a next step in the evaluation, the total number of references was divided among the project partners for an in-depth evaluation of the references for usage in the current study. References retrieved via search strings for microbiological hazards (N = 3092) were evaluated by FERA, for chemical/toxicological hazards (N = 6322) were evaluated by RIKILT, for nutritional hazards (N = 1103) and the DALY/QALY concept (N = 115) were evaluated by DTU, and for the socio-economic methodology (N=3578) were evaluated by UNEW. The references were stored into four different Endnote files, one for each project partner. Each partner evaluated the references and divided them into the three groups of relevant, maybe relevant and irrelevant references in the distributed Endnote file. Additional duplicate records were removed, and included in total: 301 duplicates for chemical hazards, 3 for DALY/QALY, 54 for nutritional hazards, 160 for microbiological hazards and 182 for socio-economic hazards. References in the Maybe relevant group were further studied and, if necessary, full texts were retrieved, to finally judge whether these references were either relevant or

irrelevant. Furthermore, full texts of the references from the relevant group in tier 1 were retrieved and further evaluated for their relevance for this study. This resulted in the following numbers of references classified in the various groups in tier 1 (based on title, abstract and keywords) and tier 2 (full text evaluation of the relevant and maybe relevant references of tier 1), see Table 1.

Table 1: Results of the literature search in the two-tier approach

Subject (per partner organization)	Tier 1: Title, abstract, keywords			Tier 2: Full text	
	Not relevant	Maybe relevant	Relevant	Not relevant	Relevant
Chemical hazards	5769	79	173	5943	101
Microbiological hazards	2601	74	257	2844	110
Nutritional hazards	979	58	12	1045	4
DALY/QALY concept	90	13	9	98	18
Socio-economic methods	3296	47	15	3366	20

The total numbers of references in tier 2 may be higher than in tier 1 due to snowballing citation, references additionally found by other partners, references obtained through internet search etc. References that were classified as relevant in tier 2 (based on full texts) were included in the Excel template as indicated in 2.3 in order to characterize the methods used.

3.2. Harmonization of terminology

The study focuses on methods that could be used for risk ranking of feed and food related hazards, in relation to human health impacts. Hazards are defined as those agents that can be present in food and feed and can affect human health. These hazards could arise from food nutrition, and/or presence of chemical hazards in feed and food, and/or presence of pathogens in feed and food which could harm human health. Regarding microbiology, the focus was thus on pathogens and not on other micro-organisms (bacteria, fungi etc.) as they do not harm human health via feed/food consumption. The same applies to chemical hazards, where we focused on chemicals that may be present in feed/food and affect human health by food consumption. Water consumption was considered to be part of food intake. The focus is on risk ranking methods for hazards related to feed and food, but the methods could also come from other areas such as natural – life sciences and social-economic sciences. In fact, these fields of application were included in the literature search to ensure the wide range of fields to which risk ranking could be applied, was included in our study.

Hazards were ranked based on the presence of the hazards in food and their severity (i.e. impact/effect on human health). In the first instance, terms for risk prioritization, risk ranking, risk analyses, and risk assessment were all used in the literature search. Though "risk assessment" and "risk analysis" imply something different from the "risk ranking" we focus upon, we also used the more general terms "risk analysis" and "risk assessment", to obtain - at first instance - the maximum possible relevant references. For example, in the example topic related to the term "parasite*", most of the references in the relevant and non-relevant categories included the term "risk analysis". There was, however, no direct reference to "risk ranking" even if there was potential to employ an approach used or developed in the particular study for use as a risk ranking method. As another example, in a small subset of references ('relevant category' for the example topic of the "consumer*" term) only one reference used the term 'risk ranking'. In other references, the terms 'Risk assessment' and 'relative risk' were used, as well as 'risk reduction' and 'risk mitigation' whereas 'Risk ranking' or 'risk prioritisation' did not appear. Nevertheless, all relevant references report on methodologies capable of comparing and ranking impacts. Risk ranking and risk prioritization and relative risks were all considered to point to risk prioritization; in this report we will use the term 'risk ranking'.

3.3. Description of risk ranking methods

Full text analyses were performed on references that described a risk ranking method in which both exposure and effect items were included. Methods that were based on only exposure or only effects were not considered. Methods described were qualitative (such as decision trees or flow charts), semi-quantitative (in which scores are attributed to exposure and effect items) or quantitative (such as risk assessments). Based on the evaluation of the methods described in the relevant references, the methods were classified, according to methodology, into: Risk Assessments (RA), ratio methods, scoring methods, Burden of Disease (Cost of illness and DALY/QALY), Willingness To Pay (WTP), Multi Criteria Decision Analysis (MCDA), Risk matrices and flow charts (including decision trees and influence diagrams), and methods based on expert judgment. All methods described estimated both exposure and effect. The way these two factors were included differs between the various methods ranging from quantitative to qualitative. Table 2 presents a summary of the numbers of references that presented a particular method category.

Table 2: Overview of method categories used for risk ranking of the various hazards

Type of hazard	Risk assessment	Comparative risk assessment	Ratio	Scoring	Cost of illness	DALY/QALY	WTP[1]	MCDA[1]	Risk Matrix	Flow chart / Decision trees	Expert judgment
Chemical	19	0	31[2]	19[3]	1[2]	9[3,4]	1[2]	13	12	13	0
Microbiological	72	0	6[2]	5[3]	9[2]	19[3]	6[2]	4	4	7	14
Nutritional	4	2	1	0	0	1[4]	0	1	0	2	0
Other	0	0	0	0	0	0	1	1	0	0	1
Sum	95	2	38	24	10	29	8	19	16	22	15

[1]WTP: Willingness to Pay; MCDA: Multi Criteria Decision Making; [2]One reference describes both chemical and microbiological hazards; [3]Three references describe both chemical and microbiological hazards, [4]One reference describes both chemical and nutritional hazards,

The following paragraphs present an overview of the results found for methods ranking chemical/toxicological hazards, microbiological hazards and nutritional related hazards. Separate sections present the results for methods used for disease burden estimation and methods from the socio-economic field that could be used to prioritize food and feed safety hazards. For each method the strengths and weaknesses are described for ranking chemical, microbiological or nutritional hazards.

3.3.1. Chemical hazards

Initial results from in depth, full text analysis of the 85 references – retrieved using search strings related to chemical hazards – are given below. These references rank chemical hazards considering human health, environmental impact, or both. Based on this evaluation, the following methods of risk ranking for chemical hazards were identified:

1. Risk Assessment

2. Risk ratio

3. Scoring methods

4. Risk matrices

5. Multi Criteria Decision Analysis (MCDA)

6. Flow charts/Decision Trees

3.3.1.1. Risk Assessment

Scope: A risk assessment for a chemical hazard aims to estimate the risk for human health associated to the presence of the chemical in various food products, and total food consumption. WHO (2009) has provided extensive guidelines regarding the principles and methods for the risk assessment of chemicals in foods thus outlining the standard, 'traditional' risk assessment, consisting of the following four steps:

1. Hazards identification: identifying which hazards may be present in the food

2. Exposure assessment: Evaluation of the exposure of an organism, system, or (sub)population to an agent (and its derivatives),

3. Hazard characterisation: The qualitative and, wherever possible, quantitative description of the inherent properties of an agent or situation having the potential to cause adverse

---

effects. This should, where possible, include a dose-response assessment and its attendant uncertainties,

4. Risk characterization: The qualitative and, wherever possible, quantitative determination, including attendant uncertainties, of the probability of occurrence of known and potential adverse effects of an agent in a given organism, system, or (sub)population, under defined exposure conditions.

The outcome of a risk assessment is an estimate of the likelihood of adverse health effects in a human population as a consequence of the exposure to the chemical hazard via food consumption. This outcome is the result of the risk characterization, which is the fourth and last step of a risk assessment, which takes into consideration the results of the three previous steps of hazard identification, hazard characterization, and exposure assessment. For threshold acting agents, population risk is characterized by comparison of the ADI (or other measures) with exposure. In this case, the likelihood of adverse health effects is notionally zero when exposure is less than the ADI. For non-threshold acting agents, population risk is the product of exposure and potency. When separate risk assessments are performed for various chemicals, those chemicals can be ranked on the basis of their respective estimated likelihoods of adverse health effects.

Application area: Risk assessments require a lot of data, and also models to fit to the data, such as to draw inferences from the data. Both data and information availability, and selection of models may give rise to uncertainties in the outcomes because of lack of necessary data, interpretation of epidemiological and toxicological data, and model uncertainties. The data needs for risk assessments are very high; available scientific data will in most instances not be enough to provide a high degree of certainty. Risk assessors have to use the best information that is available and provide insights into the uncertainty in their estimates.

Approach: Many risk assessment studies on chemical hazards have been published; about one third of all tier 1 references describe such an application study of a risk assessment for a particular chemical hazard. However, as the procedure for these risk assessments is comparable, only references describing guidelines for performing a risk assessment for chemical hazards were included, rather than studying the risk assessment applications. Note that various references indicate a risk assessment was performed, but in fact (studying full texts) either an exposure assessment or a hazard characterization was performed and not a full risk assessment.

Although general procedures are similar, risk assessments for chemical hazards in food are performed differently, e.g., using a deterministic approach as was described by e.g. (Son et al., 2013) or a probabilistic approach, as was described by e.g. (Kruizinga et al., 2008; Spanjersberg et al., 2007). The latter authors also applied a sensitivity analysis to determine the most influential factors within the risk assessment. Furthermore, different approaches were used for the exposure assessment and the hazard characterization steps.

For exposure assessment, multi-media mass balance modelling, i.e. holistic modelling, can be used which combines fate and bioaccumulation processes so as to evaluate potential exposure and risk (Arnot and Mackay, 2008; Mitchell, Arnot, et al., 2013; Wambaugh et al., 2013). Such approaches are

included in the Risk Assessment IDentification And Ranking (RAIDAR), and the UNEP-SETAC Toxicity models (USEtox) (Mitchell, Arnot, et al., 2013). Both RAIDAR and USEtox are far field - pertaining to the greater environment - ranking exposure models based on analytically calibrated methods (Arnot et al., 2006; Gouin et al., 2012; Wambaugh et al., 2013). Although these methods were applied for ranking environmental chemical hazards, the approach can also be used as input to an exposure assessment for human health risks.

Similarly, within the hazard characterization step, several methods are presented in order to rank the toxicity of the chemical hazards. For example, the toxic equivalency factor (TEF) approach indicates the relative potency of individual compounds as compared to a reference compound. The overall toxicity (or Toxic Equivalent, TEQs) is defined as the concentration of the individual compound in a mixture times the TEF, which is then used to rank the toxicity of chemical compounds (Safe, 1998). In case substances are to be ranked for which no or few data are available, the TTC approach can be used. This approach classifies compounds based on their chemical functional groups and structural features (Feigenbaum et al., in press). The approach is conservative and less accurate than applying ADI's, but can be used for compounds such as food contact surfaces with limited toxicity data.

EFSA (2011b) has published an overview of procedures for current risk assessment methods for dietary exposure of different chemical substances. They emphasize the need for not only harmonized approaches, but also for a future exploration into cumulative exposure assessments.

Strengths: In a risk assessment all scientific and technical information and data, as well as variability and uncertainties, are systematically organized. It is thus a very structured method, providing insights into what is known and the gaps in knowledge.

Weaknesses: A risk assessment for one chemical hazard will need a lot of data, knowledge and resources (manpower, money). Risk ranking of various chemical hazards in food using outcomes of individual risk assessment will take even more resources. Moreover, uncertainties related to chemical risk assessments are very high because of data limitations. Ranking of chemicals may be difficult, with large and overlapping uncertainty ranges for the risks of the different chemicals.

Perspectives for use by stakeholders: Risk assessments for chemical hazards in food are very useful to perform, amongst alternative methods, for providing insights into gaps in knowledge and issues with the greatest uncertainties. They will, however, not pose a very suitable method for risk ranking given the large amounts of data, knowledge and resources needed.

3.3.1.2.   Risk ratio

Scope: Risk ratios or quotients are derived by dividing estimates of exposure by effect. For this purpose, data are needed on the amounts of the hazard consumed (either the dose or the concentration) as well as a measure for toxicity for the hazards that are studied (e.g. ADI, TDI, RfD).

Application: The risk ratio method has been applied for ranking of a range of chemical compounds. Usually it was applied to quickly screen the risk of chemicals, mostly for pesticides (in total 13 references).

Approach: Some references derive a Hazard Index, in which the Estimated Daily Intake (EDI) is divided by the Acceptable Daily Intake (ADI), Tolerable Daily Intake (TDI) or the acute Reference Dose (RfD), see for example, the studies published by (Calliera et al., 2006; Oldenkamp et al., 2013; Sinclair et al., 2006). The Margin of Exposure (MoE) approach is another method in which exposure and effect are compared by dividing the NOAEL (No Observed Adverse Effect Level) or the BMD (BenchMark Dose) by the EDI (Bang et al., 2012; Madsen et al., 2009; Rietjens et al., 2008). The Hazard Index should be as low as possible, whereas the MoE should be as large as possible in order to obtain a low risk for human health. In general, the risk of pesticide residues for human health is ranked using the Hazard Index (e.g. (Labite and Cummins, 2012; Sinclair et al., 2006; Travisi et al., 2006; Whiteside et al., 2008), whereas the risk of carcinogenic compounds is primarily ranked using MoE (Dybing et al., 2008; Lachenmeier et al., 2012).

Strengths: This method can easily be applied once concentration data and toxicological reference values are available, and is easy to understand. A full risk assessment is not necessary, rather an estimate for both amounts of the hazard consumed and the effect of the hazard on human health.

Weaknesses: For emerging chemical hazards, such as nanomaterials, toxicological reference values are usually not available. Furthermore, concentration data are also not always easily available. It may thus be difficult to rank all hazards of interest due to data limitations.

Perspectives for use by stakeholders: The method can give a quick answer on the risk of chemical hazards for human health. It will provide a prioritized list of chemicals from low to high risk, which can for instance be used as input to establish risk based monitoring programs or to prioritize scientific research budgets.

3.3.1.3. Scoring methods

Scope: This method for scoring exposure x effect is based on semi-quantitative scoring both the exposure and the effect items, followed by their multiplication.

Application: Soring methods provide a simple risk ranking method that is often implemented to easily characterize chemical hazards for subsequent categorization into particular groups (Aylward et al., 2013; Bietlot and Kolakowski, 2012; Bu et al., 2013; Greim and Reuter, 2001; Taxell et al., 2013; van Asselt et al., 2013). It provides a screening tool that is often implemented as a pre-analysis to understand the potential risk before more in-depth risk assessments are later utilized.

Approach: When a scoring method is applied, both exposure and severity (or effect) are considered. However, endpoints for exposure and effect can vary.

Various endpoints are used to estimate exposure in risk ranking, such as chemical transformation properties (degradability, half-life), mobility/distribution (BAF, BCF, Kow etc.), release and frequency of detection, and dose administered/concentrations. There is currently no scientific consensus on which endpoints to include and how to set criteria for classifying these endpoints. Consequently, these exposure endpoints can be based on LD50, NOAEL, BMDL etc. for toxic effects.

In particular, criteria were based on BCF, BAF values, concentration levels found, or they were more qualitative using an exposure potential.

Endpoints for effect may include effects on human health with endpoints such as acute toxicity, carcinogenicity, reproductive toxicity etc. as well as effects on the environment with varying endpoints such as eco-toxicity (incorporated within Life Cycle Assessments). Once criteria are set, endpoints are classified either qualitatively, for instance, using the categories high, medium, low, as applied in, for example, (Robichaud et al., 2005), or semi-quantitatively for instance using scores from 1-3 or from 1-5, as applied in, for instance, (Penrose et al., 1994). Greim and Reuter (2001) have classified carcinogenic compounds into five categories for their impact, ranging from compounds that have a significant contribution to cancer risk to compounds which play only a minor role. Based on animal experiments and/or epidemiological studies, compounds can be classified into one of these five groups and can be scored as such.

When a classification system is established, data sources need to be found in order to assign scores. These sources can be based on literature, available data and/or expert opinion. Several methods are available for assessing toxicity varying from classifications based on physico-chemical properties (as is relevant for nanomaterials), animal experiments and/or modelling frameworks (such as QSAR to predict toxicity effects with/without including a correlation with physicochemical properties), or PBPK-models to determine the exposure-dose-response relationship (Hristozov et al., 2012). Then, scores need to be aggregated, which is mainly done by multiplying exposure and effect (see for example (Gamo et al., 2003; Juraske et al., 2007; van Asselt et al., 2013), although one study added the scores (Penrose et al., 1994).

Some references also use a weighing system to weigh the various endpoints included in the assessment (Dabrowski et al., 2014; Juraske et al., 2007; Penrose et al., 1994; Valcke et al., 2005l). A general framework for risk ranking that includes the choice of endpoints, weighing endpoints and aggregating the scores into a final risk score is depicted in Figure 1.
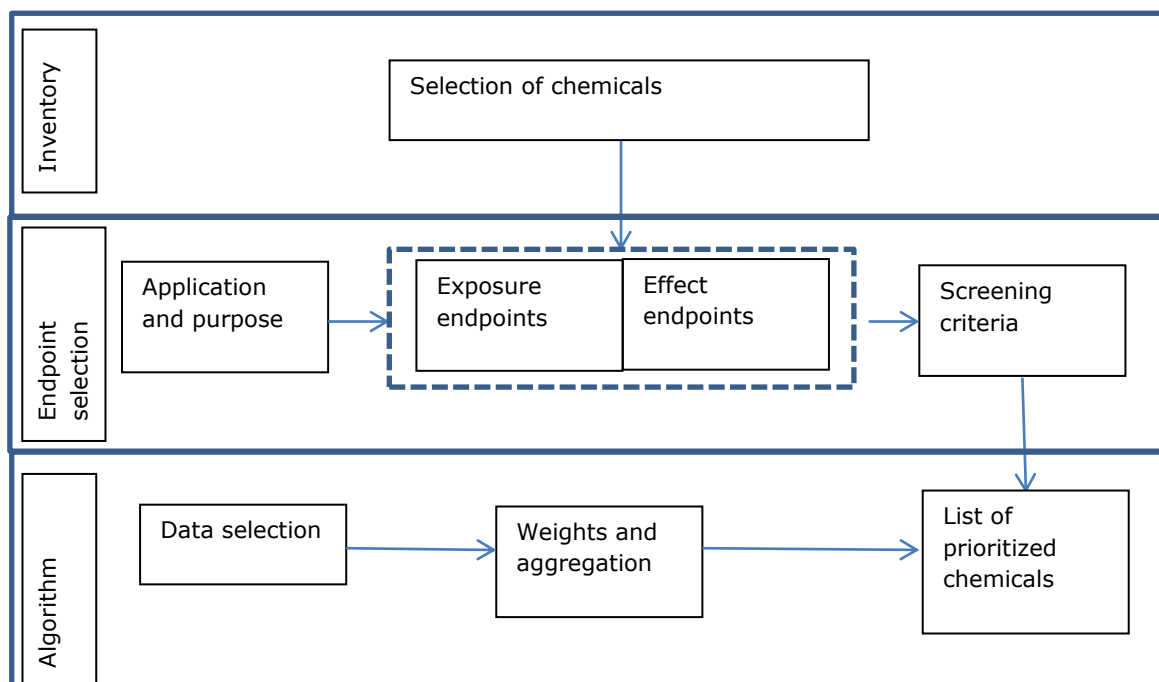
Figure 1: Framework for risk ranking of chemicals, adapted from Bu et al. (2013).

Strengths: This method is easily applicable, once scores are assigned to the model variables. Furthermore, it allows the inclusion of stakeholder perception in the scoring. The various model variables can be equally weighted, or a higher or lower importance can be assigned. The assignment of weights should be clearly documented to guarantee a transparent approach.

Weaknesses: The method is semi-quantitative and therefore less accurate than, for example, risk assessment. Furthermore, the assigning of weights relies on expert input and thus it is important to perform a thorough expert judgment study as input to the risk ranking method.

Perspectives for use by stakeholders: Stakeholders can use this method to obtain a clear overview of prioritized risks. The method has been used as input to the establishment of national monitoring programmes (VRC, 2010).

3.3.1.4.   Risk matrices

Scope: Just like the scoring methods, risk matrices also make use of scoring both exposure and effect items. The difference between scoring methods and risk matrices is that in the latter, the exposure and effect elements are not aggregated by multiplication or addition, but depicted in a risk ranking matrix with effect on the one axis and exposure on the other.

Application: This method is usually applied to chemical hazards for which limited quantitative data is available. This method has, for example, been applied for ranking of nanomaterials (O'Brien and Cummins, 2011; Sorensen et al., 2010; Zalk et al., 2009).

Likelihood	Consequences				
	Insignificant	Minor	Moderate	Major	Severe
Almost certain	M	H	H	E	E
Likely	M	M	H	H	E
Possible	L	M	M	H	E
Unlikely	L	M	M	M	H
Rare	L	L	M	M	H

Figure 1: Example of Risk matrix

Approach: In this case, both the likelihood of occurrence and the consequences of the hazard are scored into one out of several classes; see Figure 2 for an example. Classes that could be used for likelihood of occurrence could be: almost certain, likely, possible, unlikely and rare. Classes that could be used for the consequences could be: insignificant, minor, moderate, major and severe. Then, risk classes are assigned to the combinations of likelihood and consequences, for instance being L (low), M (moderate), H (high), and E (extreme), as shown in Figure 2. But, risk classification may also be based on scores. Zalk et al. (2009), for example, classified nanomaterials based on scores for probability and severity, in which scores <25 were indicated as low risk, scores between 26-50 as medium risk, scores between 51-75 as high risk, and scores >75 as very high risk. These results were then depicted in a risk matrix. Clarke et al. (2011) evaluated a variety of compounds in sewage sludge using five different criteria, using scores 1-2 or 1-3. Scores were then summed over the five criteria for each compound, to obtain a final score. Another method to visually present risks is the spider web. Ranke (2000), for example, classified various endpoints using scores from 1-4, and compared spider webs for the various compounds to obtain an indication of the most risky ones.

Strengths: This method gives a visualisation for both effect and exposure of the hazard. It provides insights into the way these two elements contribute to the overall risk of a compound. A compound may provide a high risk due to high exposure, but low severity or due to its high toxicity, but low exposure. The matrices will give more information to the risk manager as compared to other methods which result into a list of compounds with their overall risk only.

Weaknesses: This method is qualitative or semi-quantitative, and thus less accurate than methods based on concentration data and dose response relationships or toxicological reference values. Classification for consequences and likelihood may also not be fully underpinned. Furthermore, the method depends on expert input. It is thus essential to perform a thorough expert elicitation study.

Perspectives for use by stakeholders: In case stakeholders prefer a graphical representation of the risks, this method can be used to visualize both the effect and the exposure of a hazard. This facilitates discussions amongst stakeholders regarding the risks of various hazards.

3.3.1.5.   Multi-Criteria Decision Analysis

Scope: MCDA (Multi-criteria Decision Analysis) is a widely discussed and applied approach, with a long history of use in various decision contexts, for example, in nanomaterial risk assessment. It is typically used when multiple conflicting criteria need to be evaluated for making decisions.

Application: The method has been applied in many studies, for example, by Linkov and co-authors (2013) who used MCDA to evaluate various manufacturing techniques for nanoparticles. Although applied in a different field, the described methodology could also be used for ranking of food related risks to human health. MCDA allows risk assessors/managers to compare various alternatives using weights for various input factors.

Approach: MCDA is typically applied to decision making problems with multiple, often conflicting, criteria that need to be evaluated. It helps structuring and solving problems, such to lead to more informed and better decisions. MCDA is based on the identifying the different dimensions and their underlying criteria of the problems, and assigned weights or scores to the criteria and dimensions. Linkov et al. (2013) assigned scores to risk, cost and benefit dimensions such as health risk, life cycle environmental impact, financial cost, energy yield and material efficiency. Based on literature and expert judgment, either quantitative scores or qualitative scores (high, medium, low) were assigned to each criterion. Then, weights were assigned to the various dimensions, and MCDA was applied (i.e., scores were multiplied and summed across criteria to give a net score for each alternative).

Mitchel et al. (2013) presented an overview of an MCDA framework for ranking chemical hazards, considering factors such as the chemical properties (persistence, bioaccumulation, toxicity) and life cycle properties (production, consumer use, disposal) to provide an overall transparent approach (Mitchell, Pabon, et al., 2013; Sailaukhanuly et al., 2013; Sorensen et al., 2010).

Weight of Evidence (WoE) is sometimes mentioned, in which assessors, who may be stakeholders or experts, weigh the importance of various endpoints and/or data sources. MCDA is then applied to aggregate the outcomes (Critto et al., 2007; Zuin et al., 2011). Uncertainty can be included in the analysis using probabilistic distributions for the scores, and Monte-Carlo simulations to determine the likelihood for certain scenarios (Linkov et al., 2013; Sailaukhanuly et al., 2013).

Strengths: MCDA allows to consider inputs from stakeholder perception by assigning weights to the various criteria used in the analysis. Furthermore, apart from human health criteria, economic impact or other criteria that are deemed relevant can be included. This makes the method broadly applicable allowing risk assessors/managers to determine the impact of various criteria on the overall risk ranking of hazards. This method thus allows to include subjective elements that may also be important for risk managers to include in their decision making process, depending on the aim of the ranking exercise.

Weaknesses: The MCDA outcome is more difficult to communicate than more straightforward methods such as risk matrices or scoring methods as various criteria are included each having different weights. Furthermore, this method needs expert or stakeholder input in order to derive the weights for the criteria.

Perspectives for use by stakeholders: This method is very valuable in case stakeholder perception is to be included in the risk ranking as a weighing can be assigned to the various model variables. This method also allows the inclusion of factors other than effect and exposure, which makes it a very versatile tool.

3.3.1.6.   Flow charts/Decision trees

Scope: Flow charts or decision trees are based on a set of clearly defined questions or criteria. When going through these questions or criteria the chemical hazards can be classified into different categories (e.g. high, medium or low) for their risk for human health.

Application: Flow charts or decision trees can be used for various purposes. Haase et al. (2012), for example, established a decision tree for nanoparticles to determine whether a full risk assessment is required or not. In general, these methods are used to obtain a qualitative indication of the risks of chemicals.

Approach: EFSA described guidelines for classifying chemical hazards in negligible, low, medium, and high risks (EFSA, 2012d, 2012e). A general flow chart was derived for ranking residues and hazards based on the likelihood that specific residues or hazards are present in the carcass; evidence of occurrence or incorrect use in pork or poultry, the toxicological profile, and the outcome of national monitoring programmes. It is unclear how these factors were scored and aggregated to come to the overall classification.

Aschberger (2011) derived a decision tree for ranking nanomaterials based on an estimation of environmental concentration (PEC) and human exposure (occupational exposure and consumer exposure). Toxic effects were based on eco-toxicological studies to derive an indicative no-effect concentration for environmental risks, and animal studies were used to derive an indicative human no-effect level. Exposure levels were then compared to the toxicological endpoints.

Morgan (2005) used an influence diagram to determine which variables influenced both exposure and toxicity. The assessment was based on questionnaires sent among scientific experts. Once weights are assigned to these variables, such an influence diagram can be used to assess the expected value of additional information from more research in order to then fill the data gaps. This will enable risk managers to prioritize research needs.

Decision trees may also be more quantitative using threshold values to classify chemicals. Eisenberg and McKone (1998) used a Classification and Regression Tree Algorithm (CART) to specify the chemical and environmental properties. Monte Carlo simulations were used to estimate human exposure. The outcome was then analysed by CART.

Schmidt et al. (2011) utilizes a decision support system (DSS) to rank GMOs. The DSS includes all elements according to EFSA guidelines (EFSA, 2011a): compositional analysis, molecular characterisation and toxic, allergic effects. Additionally, effects on soil and the risk of protein transport was also included. The DSS is based on a decision tree and rules, indicators and baselines, and thresholds. For example, an indicator could be the level of amino acids. Baselines are then obtained from conventional plants and thresholds may, for example, be the LD50, LC50 or ADI. The DSS is incorporated in an ICT-tool and can be used to establish the risk of a GMO as well as rank the GMO within the natural variation of plants. Input is provided through literature and expert opinions (Schmidt et al., 2011). DSS may also be combined with MCDA. Critto (2007), for example, utilised a DSS system to evaluate ecological observations and eco-toxicological tests for contaminated sites and then incorporated MCDA and expert judgments into the ranking. The approach used in this paper can also be applied for ranking food safety risks.

Strengths: Flow charts/decision trees present a straightforward method with black and white questions for which only qualitative information is needed. The method can thus best be used for a quick screening of food safety hazards, of which the most relevant ones may later on be investigated in more detail.

Weaknesses: This type of method depends strongly on expert input and it is thus essential to perform a thorough exert elicitation study. Furthermore, this type of method is less transparent than the other methods presented as it is not always clear why compounds end up as high, medium or low risk. Therefore, for each compound classified based on a decision tree or flow chart, the underlying reasons for the answers should be clearly documented in order to obtain a transparent classification.

Perspectives for use by stakeholders: It is important to set up the right questions for inclusion in a decision tree based on expert judgment and scientific evidence, which may be challenging to achieve. However, once a decision tree has been drafted, it is easily applicable for stakeholders to classify hazards into high, medium and low risks.

3.3.2. Microbiological hazards

Initial results from in depth, full text analysis of the 110 references – retrieved using search strings related to microbiological hazards – are given below. These references rank microbiological hazards considering human health. The literature search revealed various methods categories for addressing relative risk ranking of microbiological hazards in food and feed (see Table 2). In this part the following methods are described in more detail:

1. Microbiological Risk Assessment,

2. Multi Criteria Decision Analysis.

MCDA often make use of expert judgments. Since, MCDA is already outlined above (section chemical hazards), particular emphasis will be given to the use of expert judgement in MCDA.

In addition to the method categories mentioned above, a number of other approaches were encountered. So, for example, flow charts, decision making trees, and risk matrices are also employed in relative risk ranking of microbiological hazards, but the examples identified were much fewer. Some interesting results can be found in Grace et al. (2012), in particular fault tree models that can be used to portray and estimate the exposure to a pathogen, Emmanuel et al. (2009) and Batz et al.(2012). The latter study is in particular quite useful to identify outcomes of illnesses and characterise the severity and relevant frequency, thus facilitating prioritisation. Finally, an interesting approach employing both forward and backward approaches to build risk matrices can be found in Stella et al. (2013).

3.3.2.1. Microbiological Risk Assessment

Scope: Microbiological Risk Assessment (MRA) aims to describe in detail the risks to the human population due to microbiological hazards in food/feed sectors (Codex Alimentarius, 2012). MRAs follow the classical scheme of risk assessments, in particular the following steps:

1. Hazard identification: identifying which pathogens may be present in the food and cause adverse health effects,

2. Exposure assessment: determining the likelihood and concentration of the pathogen in the particular food/feed at the moment of consumption,

3. Hazard characterisation: describing the possible adverse health effects of the pathogen, although dose response curves are not always possible to acquire,

4. Risk characterization: in which steps 2 and 3 are combined to estimate the risk of the pathogen for the human population describing the probability of adverse health effects under the given exposure conditions. .

Several types of MRA have been identified, with different combinations of deterministic, probabilistic (or stochastic), qualitative, semi-quantitative, and quantitative modelling. The most recent development is the Quantitative MRA, which offers the opportunity for a detailed risk characterisation, and inevitably we refer to this QMRA in the text below.

Application areas: As with the chemical area, microbiological risk assessments also require a lot of data, and models to fit to the data, to allow inferences to be drawn from the data. Both data and information available, and selection of models may give rise to uncertainties in the outcomes because of lack of data needed, including dose response and exposure data, interpretation of epidemiological data, and model uncertainties. The data needs for microbiological risk assessments are very high; available scientific data will in most instances not be enough to provide a high degree of certainty. Risk assessors have to use the best information that is available and provide insights into the uncertainty in their estimates.

Approach: The EFSA BIOHAZ panel has published its experiences gained with QMRA studies (EFSA, 2012b). It seems like a number of critical points were flagged: (a) the process where mandates are defined and distributed to Panels, (b) the selection of modelling approaches to support answering

the mandate, (c) the decisions on the criteria for data inclusion/exclusion, (d) the review of the outputs of the QMRA, and (e) the communication of the opinions to risk managers. Results of QMRA are dependent on the particular choices for models included in QMRA, for example, dose response models and growth models. An exemplar study on choices of dose response models is shown in Pouillot and Lubran (2011). Similarly QMRA are sensitive in the choice of statistical methods to accommodate for missing data. An example of the greater disagreement between two different statistical approaches when largest proportion of data is can be found in Williams and Ebel (2012). Some examples of better practices could be found in the literature items evaluated, in particular:

(a)   Charles and Ashbolt (2009) provide an example of citing model variables, data input per variable and assumptions/uncertainties tagged with each one,

(b)   Guillier et al. (2013) provide an example of a conceptual model and how uncertainty could be propagated through the models,

(c)   Hong et al. (2012) that provides an example where parameters' uncertainties are presented and discussed in a separate section in the text,

(d)   Xiao et al. (2012) provide an example of presenting clearly the equations used in the models and the data inputs for the model with a clear description of whether inputs were deterministic points or probability distributions,

Tran et al. (2013) provide an example for presenting probability maps for visualisation of data. The outcome of a risk ranking method may have an influence on how data are visualised, so for example risk assessment or MCDA results may be easier to visualise in such maps.

Strengths: QMRA offer the opportunity to address uncertainties in a more transparent way, e.g., via sensitivity analyses and/or modelling and simulation runs. QMRA could be one of the most efficient methods to estimate risk, including the relevant uncertainties.

Weaknesses: QMRA is still hampered by the lack of quantitative data, and often assumptions that generate high degrees of uncertainty have to be included (Malorny et al., 2008). Most of the QMRA studies are judged as very difficult to (a) pinpoint quickly and efficiently the assumptions and uncertainties involved in the methods employed, (b) identify quickly data and sources used, and (c) comprehend the innovative points in each study. The way that methods were employed, e.g. the choice of dose response models, appears to be a "get a tool from a drawer" exercise (see for examples in Mara et al. (2007).

Perspectives for use by stakeholders: Applied optimally, QMRA should disseminate key information regarding risk from exposure to microbial hazards. to policy makers, decision makers and the public.

3.3.2.2.   Multi Criteria Decision Analysis using expert judgment

Scope: The second most popular method for relative risk ranking of microbiological hazards appears to involve MCDA with using expert judgments. MCDA is based on decision theory and provides a framework for performing such expert elicitations, usually comprising three phases, i.e., the

formulation of the problem, the decision analysis part, and the analysis and presentation of results (Aenishaenslin et al., 2013).

Approach: Key elements of any MCDA are:

- The criteria included as these form the infrastructure of the decision analysis. The selection of criteria in MCDA is an expert elicitation exercise on its own, and it is advisable that the experts who are involved in this selection do not participate in the ranking/preferences elicitation exercises to avoid any biases.

- The experts involved, as different experts may give different opinions (due to different expertise, different knowledge, but also different ways of interpreting the same information).

- Selection of the right experts to be involved is crucial, as different experts may give different opinions (due to different expertise, different knowledge, but also different ways of interpreting the same information).

Application areas: Elicitation of expert judgment can be performed in a number of ways, depending on the type of information required. For example, the elicitation could aim to obtain numerical information like scoring, or quantitative information like ranges of an unknown parameter, or preferences (or rankings), or it could aim to elicitation of attribution. Examples of the latter can be found in Hoffman et al. (2008) and Batz et al.(2012). Elicitation of scoring is usually based on subjective, arbitrary scales. Examples can be found in Anderson et al. (2011) and Knight-Jones et al. (2010). The majority of the identified studies concerned the elicitation of some type of scoring, or elicitation of rankings. For example scoring of modules of a model that can be found in (Guillier et al., 2013; Humblet et al., 2012; Ruzante et al., 2010), or scoring of a number of criteria in the multi criteria model, as presented by (Anderson et al., 2011; Cardoen et al., 2009; Havelaar et al., 2010; Lake et al., 2010).

In an MCDA, a key issue that could differentiate approaches employed for performing the MCDA is the way that the weights of criteria are elicited. So for example criteria could be considered as equal, a way which inevitable simplifies individual experts' views. Alternatively experts disseminate weights for each MCDA criterion, therefore indicating the degree of the importance they put on each criterion in the MCDA outputs. Recently, alternative methods for performing a MCDA have been developed and employed in order to minimise the biases linked with experts' direct weighting of the MCDA criteria.. An example is found in Havelaar et al. (2010)where the risk scores were calculated via probabilistic inversion, therefore weights of criteria were not elicited but were inferred from the experts' rankings. This particular method carries high value as it allows for:

- Internal (i.e. out-of sample) validation. E.g., performed by (Flari et al., 2011; Havelaar et al., 2010),

- External validation, i.e., via a different group of experts. Such validation was performed in (Flari et al., 2011),

- Testing experts' consistency, e.g., performed by Neslo and Cooke (2011),

- Entering uncertainty distributions in each criterion for each scenario to be ranked[2],

- Aggregation of experts' views in a coherent way,

- Prediction of experts' rankings for scenarios other than the scenarios involved in the initial risk ranking exercise to build a model.

<u>Strengths</u> MCDA approaches that can be rigorously validated, e.g., methods like the one presented in Havelaar et al. (2010), offer strengths, as they allow for internal (out of sample) and external validation, and thus are highly auditable and reproducible.

<u>Weaknesses</u> Inevitably, the methods carry the weaknesses that are linked with elicitation of information from experts, i.e., need for having rigorous, auditable methods to identify experts, high demand for resources as training of experts in these methods and specialised risk analysts and modellers may be needed, need to consider how to elicit experts' own uncertainties on their views, opinions, judgments, and last but not least needs to consider possible ways to combine individual opinions without masking variability in the experts views.

<u>Perspectives for use by stakeholders</u>: Expert judgment approaches, including MCDA, can be applied in cases where crucial information is missing, and yet a decision needs to be made. Examples of such situations could refer to either long term policy making, e.g. rankings to facilitate resource allocation, rankings of research needs in a particular sector, and/or emergency responses where very rarely one has the particular data needed to answer policy making questions, and thus extrapolation (via expert judgment) and/or employment of proxy data (via expert judgment) can be very common.

3.3.3. Nutritional hazards

In total four relevant references were retrieved in relation to nutritional methodology using the search strings. No relevant reports were identified through Google searches of the grey literature. Various methods were used for ranking nutritional hazards:

- Comparative Risk Assessment Analysis (CRA),

- MCDA using generic inverse variance method

As listed in Table 2, nutritional references using risk assessment, ratio, DALY/QALY or Flow chart/Decision trees were also identified. These methods have already been described elsewhere and will not be described here.

---

[2] This is a very recent development by Dr Neslo in Utrecht University; personal communication.

3.3.3.1.   Comparative risk assessment analysis

Scope: A Comparative Risk Assessment (CRA) analysis estimates the number of deaths that would be prevented in the period of analysis if current distributions of risk factor exposure were changed to a hypothetical alternative distribution (Danaei et al., 2009; Micha et al., 2012).

Application area: Two examples of CRA analysis were identified within the nutritional hazards related literature. Danaei et al. (2009) performed a CRA analysis for establishing the preventable causes of death associated with dietary, lifestyle and metabolic risk factors in the United States. Micha et al. (2012) used a CRA framework to develop methods for assessing the global impact of specific dietary factors on chronic disease mortality.

Approach: A CRA analysis is measured in population attributable fractions (PAFs), which describe the total effects of a risk factor (direct/indirect) by reflecting the proportional reduction in deaths for each disease causally associated with the exposure that would occur if the usual exposure distribution had been reduced to the optimal minimum-risk exposure distribution. Micha et al. (2012) established a series of inputs needed for calculating PAFs, being:

a) Effect size (relative risk estimate) of the causal diet-disease relationship to quantify the diet-disease relationship for which probable or convincing evidence of a causal effect exists,

b) Optimal or theoretical minimum-risk exposure distribution (TMRED) to determine the optimal risk factor exposure distribution that is realistically attainable and is associated with the lowest possible disease risk,

c) Dietary risk factor exposure distribution in the population to determine the current average of usual exposure levels of the dietary risk factor in the population and,

d) Total number of disease-specific deaths (plus non-fatal events, when available) in the population to determine the absolute numbers of disease events caused by a certain disease in the population.

Data sources for obtaining these inputs include epidemiological studies, systematic reviews, meta-analysis, nationally representative nutrition surveys and mortality databases. Equation 1 shows how these four inputs are used to determine the PAF.

Equation 1:

$$\frac{\int\limits_{x=0}^{m} RR(x)P(x)dx - \int\limits_{x=0}^{m} RR(x)P'(x)dx}{\int\limits_{x=0}^{m} RR(x)P(x)dx}$$

Where,

x: the exposure level; m: the maximum exposure level; P(x): the usual exposure distribution in the population (by sex and age); P´(x): the minimum-risk exposure distribution (TMRED); RR(x): the relative risk of mortality or morbidity at exposure level x (by sex and age),

The number of deaths from each causally related disease outcome attributable to a risk factor can subsequently be calculated by multiplying its PAF by total deaths from that disease.

Strengths: A CRA analysis is a systematical assessment of unbiased data collected from national and international surveys as well as the published literature. It allows for consistent, comparable and quantitative assessment of the global impact of risk factors on disease by sex- and age-specific groups.

Weaknesses: A CRA analyses has an extensive need for knowledge and resources (manpower, money, data) which make it expensive to perform. A CRA analysis requires a lot of unbiased data, for instance, to establish exposure distributions or causal diet-disease relationships, which may often not be easily accessible or even available. The weights of different diseases are not considered. Moreover, uncertainties related to a CRA analysis can be high because of data limitations. Ranking of modifiable dietary risk factors may be difficult with large and overlapping uncertainty ranges for the different risk factors.

Perspectives for use by stakeholders: A CRA analysis offers a global assessment of the impact of dietary factors on disease mortality highly valuable for priority setting and policy making. However, it will not pose as a very suitable method for most risk ranking given the large amounts of data, knowledge and resources needed.

3.3.3.2.   MCDA using generic inverse variance method

Scope: A random effect model using the generic inverse variance method estimates the incidence rates of death from various risk factors including dietary related risk factors.

Application area: One principal application of the generic inverse variance method for risk ranking of nutritional hazards was identified (Umasunthar et al., 2013). These authors used the method to estimate the incidence of fatal food anaphylaxis for people with food allergy and related this to other mortality risks in the general population (e.g. death from murder, fire).

Approach: A qualitative approach where incidence rates of deaths (or another suitable endpoint) from a dietary factor is established based on a systematic review and/or meta-analysis of data from the published literature as well as from national and international surveys. The Newcastle-Ottawa quality assessment scale can be used for the quality assessment of cohort studies. For each study identified in the systematic review the number of events and total person-years of monitoring are extracted. Incidence rates are then calculated per million person-years (micromorts). The obtained incidence rates are pooled across studies using the natural logarithm of the occurrence rate and its standard error based on a random effect model using the generic inverse variance method. The results can be presented in a table but also graphically (Umasunthar et al., 2013).

Strengths: Data is obtained from a systematic review followed by a meta-analysis (if the data permits) which are both placed high in the hierarchy of evidence.

Weaknesses: High variation between studies in method of data capture, case definitions and limited information about prevalence in the populations studied may lead to heterogeneity between study results. This will obscure the results from the systematic review and meta-analysis and thus the variance rates build upon it. The amount of resources (manpower, money) put into this method is not reflected in the knowledge obtained from the method as incidence rates do not provide information on the modifiable risk factors that cause the diseases.

Perspectives for use by stakeholders: The underlying calculations are not transparent and involve expert inputs. The visual presentation of the results is easily understandable.

3.3.4. Disease burden methods

Among the references retrieved using the Disability Adjusted Life Years (DALY)/Quality Adjusted Life Years (QALY) search string, 15 publications were deemed relevant, including one paper retrieved by the socio-economic search string. In addition, three reports from the National Institute for Public Health and the Environment in the Netherlands (RIVM) have been included (Bouwknegt et al., 2013; Mangen et al., 2014; Van Lier et al., 2007).

Thirteen references ranked different chemical and microbiological food hazards, of which three references ranked both chemical and microbial hazards (Howard, 2007; Lim, 2012; Newsome, 2009). Three of the references used alternative methods of estimating clinical severity (Fosse et al., 2008a, 2008b; Newsome, 2009).

Five references presented methodology central to risk ranking and disease burden estimation (Crettaz et al., 2002; Hofstetter, 2002; Mangen et al., 2010; Mangen et al., 2014; Pennington et al., 2002).

Scope: The Disease burden methodology may be applied when the ranking of hazards is to consider the level of human disease or loss of productive capacity for the exposed population. Burden of disease estimates such as DALYs or QALYs may be used as the only parameter for risk ranking, but are often included as one of several parameters in a risk ranking model. Data for estimating the numbers of cases with the most relevant types of acute illnesses, chronic sequelae and mortality (also termed health outcomes) due to the exposure of the included hazards must be available. Different

types of hazards (chemical, microbiological or nutritional) require different types of data and modelling approaches, but after the final DALY/QALY calculations have been made, the risks estimates should be readily comparable.

Application area: Health-adjusted life years (HALYs) are nonmonetary health indices, where the actual health of an individual is compared with a perfect health situation (usually on a scale from 0 to 1) and this score then multiplied by the duration of that health state. The two most prominent HALYs are quality-adjusted life years (QALYs) that reflects the level of health (usually scoring 0 for death and 1 for perfect health) and disability adjusted life years (DALYs) that reflects the level of disability (scoring 0 for death and 1 for perfect health). A descriptive summary of the various HALYs to be used are presented in Mangen et al. (2014). The DALY method was developed at the WHO, and the Global Burden of Disease (GBD) Study is the most often referenced source of disability weights for specific disease outcomes (ww.who.int/healthinfo/global_burden_disease/metrics_daly/en/).The DALY/QALY approach has been used to rank different pathogens and chemical contaminations in the same food category, for different hazard-food category combinations or summarised and ranked for different food categories. Estimates of DALYs or QALYs have also been used to rank waterborne contaminants in lakes or water supplies as well as for ranking human risk factors in general.

Approach: When applying DALY/QALYs in risk ranking, different methodologies must usually be applied to estimate incidences of disease and probability of relevant health outcomes depending on the hazards to be evaluated (Crettaz et al., 2002; Hofstetter, 2002; Mangen et al., 2010; Mangen et al., 2014; Pennington et al., 2002). Comparisons of DALY/QALY estimates may be used for risk ranking directly, or the DALY/QALY estimated may be included in several of the other risk ranking methods such as risk assessment, MCDA, risk matrixes, flow charts/decision trees or in expert syntheses. DALY/QALY estimates in the referenced papers were included in the following three risk raking approaches:

- Risk Assessment

- Comparative Risk Assessment Analysis

- Multi Criteria Decision Analysis (MCDA)

DALY/QALY in Risk Assessment

In most cases, the methods used to estimate incidences of specific health outcomes (acute illnesses, chronic sequelae, and mortality), follow a risk assessment approach utilizing elements from the risk assessment methodology. This includes an exposure model to assess the concentration of pathogens or chemicals in foods, at several stages in the food supply system (e.g. primary production, manufacturing and processing, retail distribution) or in water supplies, and a consumption model that estimates the daily or lifetime consumption of the population at risk. The risk of disease for a specific hazard is usually estimated using dose-response relationships, and based on information from databases or literature about the probability of the different health outcomes the incidences of relevant disease outcomes can be estimated. DALYs or QALYs are then calculated based on the disability weights and duration of the health outcomes. This methodology was applied by Howard et al. (2007),

by Chen et al. (2013) using in the FDA iRISK tool (https://irisk.foodrisk.org/) and by Newsome et al. (2009), whom estimated the burden of disease using pseudo-DALYs; an alternative severity scores per individual health outcome. Alternative to the use of DALYs, the level of clinical severity can be based on the rate of hospitalisations and mortality, calculating risk scores (attributed cases x clinical severity index) to be ranked. This methodology was applied by Fosse et al. (Fosse et al., 2008a, 2008b).

For foodborne infections for which reporting is mandatory, modelling of underreporting is often used to estimate the 'true number' of infections in the population (Bouwknegt et al., 2013; Mangen et al., 2014; Van Lier et al., 2007). Food attribution modelling can be used to attribute chemical hazard or pathogen specific illnesses to specific foods, food categories or major transmission pathways. For each hazard included, experts are requested to indicate the most like proportions and an interval describing uncertainties for all foodborne cases (Lake et al., 2010) or proportions related to specific food categories (Batz et al., 2012; Bouwknegt et al., 2013). Food attributable proportions may also be based on analysis of foodborne outbreaks (Batz et al., 2012; Bouwknegt et al., 2013) and several other methods such as microbial fingerprinting, retrospective case-control studies or exposure assesments - methodology and choice of metrics is discussed in (Hofstetter, 2002; Mangen et al., 2010; Mangen et al., 2014).

A method for cumulative risk assessment of waterborne contaminant is presented where; toxicity measures are estimated for non-cancer and cancer effects separately, based on RfDs and Cancer Slope Factors, and in one case used for a life cycle assessment (Crettaz et al., 2002; Pennington et al., 2002); DALYs per case are estimated for non-fatal disease outcomes by multiplying DALY severity scores with the expected lifespan (80 years) and for fatal disease outcomes (cancer) years lost compared to the expected lifespan is estimated. The proportion and incidence of individuals at risk (toxicity x exposure x population at risk) is estimated and mean individual total DALY values are calculated by summing individual risk (DALYs) for each compound, and for the whole population at risk (Crawford-Brown, 2012; Etchie et al., 2013).

DALY/QALY in Comparative Risk Assessment Analysis

An approach estimating the global or regional burden of food risk of chronic diseases are measured in population attributable fractions (PAFs). The PAFs reflect the proportional reduction in deaths for each disease causally associated with the exposure that would occur if the usual exposure had been reduced to the optimal minimum-risk exposure.

Lim et al. (2012) estimate the Global disease burden attributable to risk factors by: selection of risk–outcome pairs (such as zinc deficiency and intestinal infectious disease) to be included in the analysis; estimation of distributions of exposure to each risk factor in the population; estimation of etiological effect sizes, often relative risk per unit of exposure for each risk–outcome pair; choice of an optimal minimum-risk exposure distribution; and computation of burden attributable to each risk factor, including uncertainty from all sources. This reference is part of the WHO Global Burden of Disease Study 2010.

DALY/QALY in Multi Criteria Decision Analysis (MCDA)

An approach, scoring five factors for selected hazard-food combinations including public health impact (DALYs and cost of illness), market impact on the domestic market (Value in $ of volume at retail and from export minus value from import), as well as Likert-type scales to capture consumer perception and acceptance of risk and social sensitivity to impacts on vulnerable consumer groups and industries (scoring from 1 to 0). Risk ranking is facilitated through the use of MCDA: first the hazard-food combination for each of the factors are ranked separately and then using the rank as scores to be summarised to a final score. This methodology was applied by Ruzante et al. (2010); see further description in the section on socio-economic ranking methods.

Strengths: The absolute strength of the DALY/QALY methodologies is that it readily allows comparisons between very different types of hazards, not only food related hazards but all types of human risk behaviour over time and geographical regions as presented by the Global Burden of Disease Study 2010 (Lim, 2012) and ECDCs initiative for developing methodologies for measuring current and future burden of communicable diseases (Mangen et al., 2014). Very different methods were applied for estimating the incidences of relevant health outcomes; however the following calculations of a common health index enabled a simple risk ranking based on the DALYs.

There are also several DALY calculating tools for public accesses.

- The Burden of Communicable Diseases in Europe project at ECDC has developed the BCoDE tool (http://www.ecdc.europa.eu/en/healthtopics/burden_of_communicable_diseases/project/Pages/project.aspx) that may be down loaded to a PC.

- FDA iRISK tool (https://irisk.foodrisk.org/) is a web based programme developed in the US by the Food and Drug Administration Centre for Food Safety and Applied Nutrition (FDA/CFSAN), Joint Institute for Food Safety and Applied Nutrition (JIFSAN) and Risk Sciences International (RSI).

- The WHO is currently developing the National Burden of Disease (NBD) toolkit http://www.who.int/healthinfo/tools_data_analysis/en/) that is being made available to selected research groups in Member States for experimental use.

- Also a DALY calculator for integration in the statistical programme R has been developed (http://users.ugent.be/~bdvleess/DALYcalculator/) by the University of Gent and National Institute for Public Health and the Environment in the Netherlands (RIVM).

These tools have integrated the disability weights, and depending on the programme, information about estimated or reported incidences of the different disease outcomes in the population due to one or several hazards may be entered.

Weaknesses: The concept of health adjusted life-years can be difficult to communicate, compared to estimates of monetary cost-of-illness, hospitalisation- and/or deaths rates. DALYs or QALYs do not

capture other aspects of diseases, such as emotional effects on friends and families and social or cultural difficulties.

There is a risk of over-interpretation of the relative differences, as the estimated DALY or QALY values seems as relatively precise quantitative estimates. However, DALYs and QALYs are semi-quantitative estimates based on disability scoring, and the accuracy of are highly dependent of the quality of input data and the risk assessment models used for estimating the incidences of relevant health outcomes.

A general methodological weakness is inadequate evidence to estimate the incidences of chronic disability, especially in cases with few or none symptoms during the acute phase of an infection. Another weakness is that the concept of DALYs assumes a continuum from good health to disease, disability, and death independent of time – a concept not universally accepted. It can be argued the reductions in quality of life may be reduced over time for long-term disabled persons. Also, some burden of disease studies use age-dependent severity weighting that reflect the different age specific dependencies of the community and the importance for community productivity, meaning one year lost for a 70 year-old is not as important as one year lost for a 30 year-old.

Perspectives for use by stakeholders: Tools are readily available for calculating DALYs for a range of infectious diseases including foodborne zoonoses in the EU (BCoDE tool from ECDC). If risk assessments or models for under- estimation of reported cases are available, the resources needed to estimate DALYs are moderate. However, development of risk assessment models to estimate the number of diseased individuals can in some instances be very time-consuming.

DALY or QALY estimates can be viewed as an economic measure of human productive capacity, enabling ranking the 'community production losses' related to the included hazards. However, due to the differences in methodology, comparability among studies must be considered if DALY/QALY estimates from different studies are to be used in risk ranking.

For monitoring purposes, risk ranking models estimating DALYs/QALYs can be constructed so that yearly input of surveillance and population data can be entered, as done for the food borne pathogens in the Netherlands (Bouwknegt et al., 2013).

3.3.5. Socio-economic methodology

A total of 20 relevant references were identified in relation to socio-economic methodology. Twelve were retrieved using the search strings; a further four journal articles were identified from citations in these papers, and four relevant reports were identified through Google searches of the grey literature. The main methods identified are summarized below and are characterised as either:

- Expert judgment

- Economic analysis (Costs of Illness, Willingness to Pay)

MCDA (involving integration of socio-economic criteria with natural scientific information) with their focus on human behaviour, the only likely relationship with animal studies might be if data from animal studies were used in estimating disease incidence, (for example dose-response estimates) for use in Cost of Illness calculations, or if participants in expert workshops referred to such studies.

3.3.5.1. Expert judgment

Scope: Expert judgement-based methods elicit rankings from citizens, stakeholders or other experts, and have the potential to produce a systematic and transparent ranking of risks. A variety of methods exist, for application in workshops or surveys, which may be characterised by the flows of information which take place between participants and researchers (Rowe and Frewer, 2005). There may be a one-way flow of information from experts (or other stakeholders) to researchers which aims to capture participants' existing knowledge and experience. Alternatively there may be a two-way flow, whereby participants are provided with detailed scientific and socio-economic information on which to base their deliberations and ranking, which is finally communicated to the researchers. Formal semi-quantitative techniques exist to combine divergent data sources, for example multi-criteria decision analysis (MCDA), which is described in a later sub-section (and also in Sections 3.3.2 and 3.3.4), and the Carnegie-Mellon approach. In the latter approaches, the judgement of stakeholders is used to rank risks and to allocate weights to the different attributes to produce a multi-attribute ranking.

Application area: Three principal applications of Judgement-based risk ranking were identified: a) achieving a ranking when there are data gaps, b) reconciling the diverse information streams and considerations encountered in multi-attribute problems, and c) incorporating societal values (e.g. (Moffet, 1996). Comparing risks is 'a complex mix of science and value judgements', and value judgements are implicitly or explicitly made in several areas in quantitative risk ranking including problem framing (e.g. the range of issues to be covered, selecting evaluation criteria), and in issues concerning incomplete data (e.g. developing assumptions) (Moffet, 1996). Moreover, the inclusion of public perceptions, priorities and values may cause a different ranking to be reached compared to that derived from scientific criteria alone. These might reflect public concerns such as whether the distribution of costs and benefits is equitable, the characteristics of the people likely to be affected (e.g. children or elderly people), whether exposure to the risk is voluntary or involuntary, and whether there is 'dread' or fear of a catastrophic impact. It is legitimate to incorporate these, as it is the public who bear the food safety risks and who provide the resources for dealing with them (DeKay et al., 2005).

Approaches: Approaches vary according to whether they involve experts or lay people, the amount of technical information about risks and impacts that is provided to assist study participants, whether the approach is qualitative or semi-quantitative, and whether or not the process involves deliberation among participants. The current study identified four approaches:

- Expert elicitation, defined as a set of formal research methods used to characterize uncertainty about scientific knowledge and to provide alternative parameter estimates when there are meaningful gaps in available data (Batz et al., 2012; Bouwknegt et al., 2013). Workshops and the Classical Delphi method are commonly used approaches to expert elicitation. The Delphi technique relies on the administration of two or more 'rounds' of iterated mail questionnaires to

participants possessing expert knowledge and experience (Van der Fels-Klerx et al., 2002). It provides feedback of the results and comments from previous rounds, thereby simulating some of the interaction of face-to-face meetings. This approach is semi-quantitative. In a comparative study using both CoI and the DALY approach, Lake et al. (2010) used a Delphi study to estimate food-attributable proportions for a range of diseases.

- Survey based on existing knowledge of lay or expert participants (i.e. minimal technical communication during the study). Schwarzinger et al. (2010) produced a final ranking of three health problems using a semi-quantitative approach in which pair-wise comparisons of citizens' risk rankings were made. Harrington (1994) ranked occupational health problems as identified by specialist medical practitioners in a further application of a Delphi study.

- Ranking achieved through deliberation only, or deliberation with supporting technical information (e.g. focus group or workshop). Although the ranking process may be restricted to a panel of experts considering scientific data only (e.g. US EPA's Unfinished Business project, 1987 which elicited expert opinions to rank environmental risks), there is also the possibility to involve lay people and thus capture societal values.

- Carnegie-Mellon approach which was specifically developed as a standardised procedure by which several risks could be ranked, and involves the elicitation of the explicit preferences of lay groups (DeKay et al., 2005). There has been little focus on this approach in food risk prioritization, although the procedure will be broadly the same as that proposed for environmental health studies. The basic procedure requires expert technical inputs to define and categorize the risks to be ranked, to select attributes by which the risks are characterised, and to prepare risk summary sheets to assist deliberations on each risk (Florig et al., 2001). Attributes might include inter alia disease impacts, probability of impacts, uncertainty, ability to control the risk, and degree of scientific understanding (DeKay et al., 2005). Risk summary sheets provide concise information about each attribute. Ranking of risks is performed by lay people (not experts) in a workshop setting according to their levels of concern about the risks, having considered the information provided on the risk summary sheets. By this means, social value judgements, made by members of the public or their representatives, are incorporated into the ranking. They may rank the risks holistically, or attribute-by attribute for each risk, or by multi-attribute ranking with unequal weights. Individual and group rankings can also be obtained. If used, weights for each attribute are obtained from each participant and reflect social value judgements. The procedure used for weighting is much simpler than that typically used in MCDA. The Carnegie-Mellon approach is also distinguished from MCDA studies by the amount of deliberation that occurs, the production of holistic rankings and the production of a 'thick description' of the results to help decision makers interpret the results (DeKay et al., 2005).

Strengths: Judgement-based methods provide additional information to that of technical assessments, for example where a problem is poorly understood, or technical data are incomplete. The outputs commonly include a narrative component which can make explicit the interpretations and assumptions which underlie the final ranking, as well as identifying the difficulties and uncertainties which

determine its limitations. They also provide a means of engaging the general public in evaluative and decision-making processes and of incorporating societal preferences for different alternatives. The Carnegie-Mellon approach is a systematic and transparent process which combines scientific and risk perception factors. It offers the potential for widespread use (transferability) of 'ready-to-use' information sheets relating to a particular problem. It uses simple calculations and produces various rankings and rich description. There is potential to build up a database of rankings which can be applied to novel risks as a desk study (DeKay et al., 2005). It appears to be a transferable tool, requiring the capability to run focus groups and workshops, and perform relatively simple calculations.

Weaknesses: Judgement-based methods require very careful design if they are to provide valid outcomes. Biases are introduced by a number of means including: inappropriate selection of the participants; the framing of the problem(s) for consideration; the way the process is conducted such that the whole range of opinions may not be elicited and recorded, and the content of the technical information that is presented to participants (e.g. bias, comprehensibility, acknowledgment of its limitations). Due to this meticulous preparation they are often resource intensive. Furthermore a qualitative analysis of data (if required) makes heavy time demands both in the transcription of audio recordings and their subsequent (thematic) analysis.

Perspectives for use by stakeholders: Unless judgement-based methods are planned and executed well there is a danger that they will be biased and unreliable. Depending on the specific method, the output may be a simple ranking, but could also be a lengthy narrative which, though having explanatory power, requires lengthy consideration. These methods can provide input in cases where crucial data are missing, and a decision needs to be made. They provide a means of incorporating societal values into risk ranking.

3.3.5.2. Economic Analysis

With regard to economic analysis, in which the monetary impacts of feed/food related hazards are estimated, the search identified two methodologies:

- Cost of Illness

- Stated Preference techniques

Cost of Illness method

Scope: The Cost of Illness methodology implies calculating costs related to disease and death in society due to feed/food related hazards. It can be applied wherever there are quantitative data relating to the impact of disease (severity and duration; mortality) and sufficient cost data for calculating resultant treatment costs and losses of income. Subject to data availability it is possible to compare large numbers of food risks.

<u>Application area</u>: This approach can be applied for comparing diseases (Gadiel, 2010), for food-disease combinations (Batz et al., 2011), and for supply chain analysis of a single food-disease combination (Miller et al., 2005).

<u>Approach:</u> The starting point of this quantitative method is the construction of a separate disease outcome tree (or equivalent) for each illness under consideration. This will show the numbers (and proportions) of the affected population who experience each type of impact, defined as the disease severity class. Figure 3 shows a relatively simple disease outcome tree. A critical point is whether the tree – and the analysis it supports - is restricted to acute effects, or whether long-term effects (sequelae) are also included. This will be particularly important for diseases for which some affected individuals will experience life-long disease, or where medical problems may be latent for a period (e.g. toxoplasmosis).
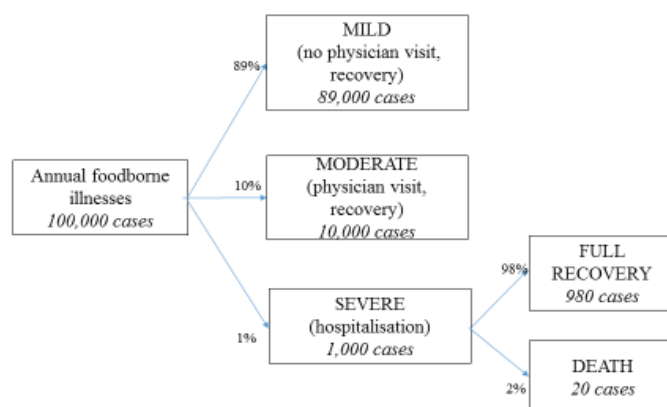
Figure 2: Disease outcome tree for a hypothetical pathogen. Reproduced from Batz et al. (2011) with the author's permission.

If possible, the disease outcome tree is populated directly from existing data sources. However, data for disease incidence and attribution to a specific food source is often incomplete. The problems with inadequate or missing data are sometimes overcome by expert elicitation of (ranges of) parameter values (e.g., (Batz et al., 2012; Golan et al., 2005)). To address uncertainty caused by inadequate data, sensitivity analysis can be used to explore the effect of varying parameter values or model assumptions (e.g (Batz et al., 2011). Alternatively, frequency distributions can be used in Monte Carlo or stochastic simulation models (Kemmeren et al., 2006; Lake et al., 2010).

The costs incurred at each state are calculated. There is some variation between studies in what costs are included, but they are generally categorised as follows:

- Direct health costs (e.g., cost of drugs, medical consultations, tests, hospitalisation), which may be further divided into acute illness and sequelae (Lake et al., 2010),

- Indirect health costs (e.g., travel to hospital),

- Indirect non-health costs, which is the value of production lost to society due to temporary absence from work of the sick person, and of the person caring for them. One of two different methods is selected for calculating lost productivity, based on assumptions about the prevailing labour market. The *friction cost* method assumes that a sick, invalid or dead worker will be replaced by an unemployed worker. The *human capital* approach assumes the vacancy (temporary or permanent) will not be filled and, therefore, considers the potential loss of production that occurs from disease onset/premature death until retirement age (Mangen et al., 2007). An alternative way of incorporating *death* is to employ a coefficient for the cost of death known as the Value of a Statistical Life (Batz et al., 2011).

Cost of Illness (CoI) analysis is limited as it includes only some of the costs which are incurred. Buzby et al. (1996) provide a comprehensive list of societal costs of foodborne illness which includes items such as industry and regulatory costs and the non-financial economic costs borne by individuals such as anxiety, pain and lost leisure time. CoI studies generally make use of discounting by which the value of earnings and payments incurred in the future are expressed in terms of their present value. They are expressed as a given amount of money invested today at a given interest rate (or discount rate) (Crutchfield et al., 1999). By definition, discounting does not apply to the costs of health effects whose duration is shorter than 1 year, whereas other end-points, such as life-long disabilities, are strongly affected by discounting. Hence, the effect of discounting will differ per pathogen (Kemmeren et al., 2006).

Strengths: The CoI method employs readily available and reliable data (Buzby et al., 1996) and the calculations are transparent and relatively simple. The resource efficiency is expected to be relatively high provided adequate data are available to produce the disease outcome trees, and computer models can be adapted for new data or additional food risks. Expert opinion may be needed to impute values that are missing. Tools for estimating CoI exist such as the on-line Foodborne Illness Cost Calculator of the USDA Economic Research Service (ERS).

The same disease incidence data are used in QALY and DALY calculations so it is relatively efficient to produce both sets of rankings at the same time and they are somewhat complementary. A combined risk ranking can also be produced. However, CoI rankings may be different from those obtained from QALY loss. For example, chronic pain that diminishes quality of life will appear in QALY calculations, but will be omitted from CoI calculations unless it causes loss of earnings or medical costs.

Weaknesses: A CoI ranking diverges from most measures of disease severity or social welfare (Golan et al., 2005) because CoI estimates are restricted to market goods. Therefore, apart from medical costs, they ignore non-workers, and don't address perceived quality of life including factors such as pain and stress (Golan et al., 2005). Taken literally, human capital estimates of foregone earnings suggest that little value is placed on reducing risk to the elderly, and rather low values are attached to reducing risk to children as their future earnings are discounted to present values (Buzby et al., 1996). The CoI method does not discriminate between the characteristics of risks, for example, whether they are voluntary or involuntary.

A further important weakness relates to the lack of accurate public health and attribution data, which is the biggest cause of uncertainty in CoI estimates. In the absence of geographically specific data, coefficients are sometimes transferred from one setting to another (e.g., Lake et al (2010) apply Dutch duration of illness data to estimate CoI in New Zealand). Results are dependent on the assumptions made *inter alia* about medical outcomes and the prevailing labour market. In theory, at least the choice of discount rate can alter the final ranking, as it will alter the discrimination between diseases with acute short-lived symptoms and those that incur costs later on (Kemmeren et al., 2006). Gadiel (2010) shows the impact of different discount rates on a range of diseases, although the final ranking is not changed. Applying a discount rate does not alter the present value of salmonellosis and norovirus (because most effects are short-lived and costs are incurred in Year 1), whereas shiga toxin-producing Escherichia coli (STEC) and campylobacteriosis, which incur future costs due to rare but severe long term consequences, are sensitive to the application of discounting. Care must be taken in comparing studies as the components of the calculations and assumptions may be different.

Perspectives for use by stake-holders: This is a well-tried technique with well-understood limitations relating to missing data and failure to adequately include non-working members of society and quality of life impacts. Large numbers of risks can be ranked. The process appears highly transparent, but it should be remembered that the cost coefficients and incidence data may be derived from inadequate data, so sensitivity analysis may be advisable. There is the prospect of updating the CoI estimates as new or better data become available. Due to non-standardisation of technique (e.g. different components, and assumptions), comparability between studies is awkward.

Willingness to Pay (Stated preference methods)

Scope: Stated preference methods could be used to elicit the preferences of individuals (citizens and households) for reducing the risk from a range of food-related diseases. When aggregated they show society's preferences. These methods take into account the concerns and perceptions of society, and consequently the ranking produced may be different from that produced by experts on technical grounds alone. Examples of the sort of questions which might be asked in a disease-based study include:

- *How much are you willing to pay (WTP) per month to eliminate all salmonella?*

- *How much are you willing to pay to reduce the risk of infection of listeriosis by 50%?*

The question could ask which alternative is preferred:

- *Would you prefer to pay a large amount for a small reduction in salmonella risk or a small amount for a larger reduction in listeriosis risk?*

The majority of published studies do not align well with the purpose of the current study because they are based round the ranking of remediation options. Such an approach is well suited to benefit-cost analysis since the aggregated value of benefits to individuals can be compared to the costs of introducing regulation.

Application area: There is a relatively long history of the use of stated preference techniques for valuing non-market goods in the analysis of environmental problems. So far, their application in ranking food risks is limited and largely confined to valuing individual disease reduction measures or comparing alternative risk management options within single food-disease problem, see for example Mørkbak & Nordström (2009) and Miller et al. (2005). Golan et al (2005) concluded that, as yet, there is not a coherent set of guidelines for conducting such studies, making comparability between studies difficult. In theory these methods could be used to rank diseases, disease-food combinations, or stages in supply chains. However, it is a complicated technique to use, which might explain the lack of use for ranking more than a small number of alternatives.

Approach: When markets exist, costs of illness (and benefits of avoidance) can be directly observed, and this is the basis of CoI analysis. However, the CoI approach ignores the value individuals place on other factors for which no markets exist such as (not) experiencing pain, lack of security and avoiding unpleasant diseases. The costs and benefits of these non-market factors, together with market factors, can be estimated using willingness to pay or willingness to accept (compensation) (WTP or WTA), a group of methods whereby a simulated market is constructed and monetary values are derived from hypothetical questions. The methods include *stated preference* techniques (*contingent valuation* and *discrete choice experiments*) and averting behaviour or preventative expenditure, which is the cost of preventing illness (e.g. paying more to drink bottled water; see (McSpirit and Reid, 2010). Stated preference methods are also able to include the value of lost health in people who are not in the labour force (e.g. retired) who are excluded from CoI calculations. Revealed preference methods also exist, for example simulated auctions.

WTP rests on the observation that people make trade-offs between health and other goods and services. The approach elicits the resources an individual is willing to give up for a reduction in the probability of encountering a hazard that will compromise their health (Golan et al., 2005). The method captures the preferences of individuals and society. For example, a skiing accident and a pathogen-borne disease might have the same probability of occurrence, but people might be willing to pay more to avoid one risk as compared to the other. This could reflect concerns such as dread (fear of catastrophic consequences), whether the risk is voluntary or involuntary, whether the illness disproportionately affects children etc. An example of contingent valuation technique is supplied by Hammitt and Haninger (2007) in which participants are asked to bid the amount they would be willing to pay to reduce the probability of food-borne illness.

Goods may be defined as bundles of attributes (characteristics) called consumption bundles from which consumers obtain utility (satisfaction). Different goods are distinguished by the specific levels of each attribute which they contain. Consumers can rank different consumption bundles in order of preference. Choice experiments (CE) use highly structured questionnaires to estimate the utility gained by the individual from each attribute or characteristic of a 'good'. In the case of food risk ranking the 'good' might be a food-related risk, defined by component attributes related to health impacts (such as infection rate, severity of illness, length of life), and the price per portion. Different goods (disease-food combinations), though defined by this common set of attributes, differ from each other in the levels of each attribute they possess (for example high, medium or low level of severity). Thus we could conceive of a 'shopping bag' type of experiment in which participants choose between different foods, different levels of potential health impacts, and the different prices paid for each option. Mørkback and Nordstrum (2009) conducted a choice experiment to elicit WTP for campylobacter-free chicken as compared to the alternatives, non-labelled chicken and outdoor-reared chicken; in other words, the WTP for higher food safety compared to the current level. Mørkbak et al. (2011) applied the technique to compare three risk remediation strategies in pork production.

Uncertainty in WTP studies arises from the heterogeneous preferences of consumers (survey respondents). Consequently reports generally show the central estimate for marginal WTP (which is the WTP for a specific incremental change in a particular food risk), standard deviation, and 95% confidence intervals for the estimated parameter value (e.g. (Mørkbak et al., 2011).

Strengths: WTP is generally viewed as the most complete and correct economic welfare measure of the benefits of food safety policies. This is because, like CoI, WTP includes the cost of treatment and lost productivity but also (unlike CoI) changes in consumer welfare such as pain, distress and inconvenience (Hoffmann, 2010). Both individual and societal WTP can be calculated. A useful feature is that stated preferences may be linked to participant profile (by means of regression analysis) revealing which societal groups (by age, background etc.) ranks a particular risk most highly (e.g., (Haninger and Hammitt, 2011)). The benefits of decreasing a particular risk are expressed as a monetary unit, for example the mean amount in Euros that an individual is willing to pay to obtain a 10% reduction in the probability of contracting *salmonella*. The aggregated value of benefits (or societal WTP) can be compared with the costs of remediation since both costs and benefits are expressed in monetary units. Though beyond the scope of the present study, these values may be used in cost-benefit analysis, cost-efficiency analysis, and Value-of-information approaches. It will be important to make sure such data are compatible, for example by relating to the same time period and the same marginal change in risk.

Weaknesses: It is argued that WTP reflects the ability to pay, and implicitly assumes that the existing distribution of resources in society is acceptable (Golan et al., 2005). However, because WTP studies can produce results segmented by sub-population, they may draw attention to unequal distributional impacts which should be considered in policy making. It is a difficult technique which is prone to errors and bias unless conducted meticulously. Experience so far has been in comparing only 2 to 4 alternatives. It may be possible to elicit mean WTP for a larger number of risks, but the scope of choice experiments may be limited by the capacity of participants to choose between a large number of choice sets encompassing many attributes.

Perspectives for use by stakeholders. These techniques provide a means of incorporating societal preferences in ranking and decision making. However, experience in the context of food risk as yet is only modest, and there is scope to develop techniques still further. (For example the 'chaining' technique might make possible an ordinal ranking of greater numbers of risks. See Chilton et al. (2002). These studies are complicated to conduct and will generally require primary data collection.

3.3.5.3.   Multi-criteria decision analysis (MCDA)

Scope. Decision-making for food and health issues is problematic as the problems are multi-dimensional and, because different people have different sets of priorities, there is no single right answer. Moreover a variety of metrics are used to measure the various dimensions. MCDA provides a fairly transparent means of identifying the salient parameters of a problem (technical information, uncertainty and different stakeholder preferences) including both quantitative and qualitative data, and integrating large amounts of complex information to allow for comparison of different risks on a common basis.

Application: To any range of problems which can be defined in terms of a common set of criteria. In recognition that the scientifically 'best' solution may be inadequate in terms of acceptability to society, or sub-optimal in terms of allocating resources, stakeholder methods are sometimes used to capture the preferences of consumers or citizens. Hence stakeholder engagement can feature in MCDA especially when politically acceptable solutions are to be defined.

Approach: MCDA is a semi-quantitative method in which a range of different criteria are identified against which each problem is assessed. Participants, either experts (e.g., (FAO and WHO, 2012), stakeholders or lay people, are supplied with technical information in relation to each risk criterion to assist their deliberations. For each risk under consideration, participants give each criterion either a numerical score or an ordinal ranking such as 'high', medium' and 'low'. The weighted scores are then combined to produce a single score for each problem, permitting scores to be ranked. The relative importance of the different weights varies between individuals, reflecting their value structure, but an 'average' set of weights can be produced for a group.

Ruzante et al. (2010) examined the risk associated with different food-pathogen combinations in relation to 6 summary measures (or factors): public health (DALY and COI), market impact, consumer acceptance and perception of risk, and social sensitivity (consumer and firm scores). Ranking was achieved by making pair-wise comparisons of alternative pathogen-food combinations on each factor enabling different criteria scales to be used without influencing the ranking. Another approach is to assign scores separately for each criterion and then express them in a common metric such as a utility scale. These can be aggregated into a single criterion (aggregate utility) for each risk that is assessed (Bonano et al., 2000; Ellis and Garelick, 2008). Ruzante et al. (2010) demonstrate that the relative importance (weighting) assigned to each criterion can substantially influence the final ranking obtained.

Strengths: MCDA explicitly and systematically combines a range of factors. Differences in participants' preferences relating to the importance of different factors are incorporated by means of weightings, either at individual or group level. Best judgement can be used when data are missing. It is

important that weightings are obtained from lay participants as the outcome will not have societal support unless care is taken to ensure that it is not just experts' preference structures that are used in deriving weightings or in selection of criteria. Once developed the method can be transferred to different expert or citizen groups.

Weaknesses: The development of appropriate information is a complicated process involving value judgements and assumptions. Scepticism has been expressed about whether it is possible to combine widely different concepts in a single metric (e.g. (Moffet, 1996). The outcome is very susceptible to framing and design, and also to the weightings which are calculated. Transparency is sometimes obscured by complicated calculations.

Perspectives for use by stakeholders: Systematic method which produces a single number for ranking. However the underlying calculations can be difficult for the non-expert to grasp.

3.4. Overall framework for selecting ranking methods

3.4.1. Design of overarching decision framework

Based on the results from the critical evaluation of the different risk ranking methods and their categories, an overall conceptual framework for risk ranking has been developed that allows for selecting the most appropriate risk ranking method. As the selection of methods depends on various parallel aspects, it was not possible to derive a decision tree. Therefore, the umbrella framework has the format of an interactive decision table for selection of appropriate risk ranking methods, based on the characteristics of the case under consideration. The decision table has been designed as an Excel tool using the main characteristics of the different categories of methods. Using this decision tool, given case characteristics and requirements, the most appropriate method(s) can be selected. This interactive framework will provide a structured guide for selection of the appropriate method(s) that could be used for risk ranking of feed and food related hazards, considering the impact on human health. This framework is presented in Table 3 and has been developed as an interactive tool in Excel in order to allow a quick evaluation of available methods.

In order to apply the decision tool, the risk manager prerequisites should be clear as well as the data availability for performing the risk ranking. For this purpose, questions are derived that allow a distinction between methods. Questions regarding risk manager prerequisites should be answered by the risk manager/assessor at the start of a risk ranking exercise. Examples are the amount of time and money available for performing a risk ranking as well as questions on what elements to include in the risk ranking (for example economic impact and/or stakeholder perception). The risk assessor is responsible for obtaining data in order to perform the risk ranking and he/she should thus answer the questions related to data availability (such as the availability of toxicological reference values or consumption data). All derived questions are indicated in Table 3. Based on the answers to the questions, one or more methods may be applicable for performing a risk ranking.

Table 3: Overall conceptual framework, or decision table, for selection of the most appropriate method(s) for risk ranking of feed and food related hazard, considering human health impacts.

Question	Answer	Risk Assessment	Comparative Risk Assessment	Ratio (Exposure/ Effect)	Scoring method	Cost of Illness	DALY/QALY	WTP	MCDA	Risk Matrix	Flow charts /Decision trees	Expert judgment
RISK MANAGER PREREQUISITIVES (Several answers per question acceptable)												
Amount of person hours resources available?	High	1	1	1	1		1	1	1	1	1	1
	Moderate			1	1	1	1		1	1	1	1
	Low									1	1	1
Level of output needed?	Quantitative	1	1			1	1	1				
	Semi-quantitative			1	1	1	1	1	1	1		
	Qualitative									1	1	1
Is it important that the methodology can be easily explained to stakeholders (laymen)?	Yes			1	1	1			1	1		1
	No	1	1	1	1	1	1	1	1	1	1	1
Should stakeholder perception be included?	Yes				1				1	1		1	1
	No	1	1	1	1	1	1		1	1	1	1
Should uncertainty be included?	Yes	1	1	1	1	1	1	1	1			1
	No	1	1	1	1	1	1	1	1	1	1	1
Should weights be included for the various risk ranking criteria?	Yes				1				1			1
	No	1	1	1	1	1	1	1	1	1	1	1
Should human incidences be included?	Yes	1	1			1	1	1	1			1
	No	1	1	1	1				1	1	1	1
Should economic impact be included?	Yes					1		1	1		1	1
	No	1	1	1	1		1		1	1	1	1
Preferred method of communication	Graph	1	1			1	1	1		1		
	Tabulation	1	1	1	1	1	1	1	1			1
	Decision Tree										1	

Question	Answer	Risk Assessment	Comparative Risk Assessment	Ratio (Exposure/ Effect)	Scoring method	Cost of Illness	DALY/QALY	WTP	MCDA	Risk Matrix	Flow charts /Decision trees	Expert judgment
DATA AVAILABILITY (only one answer per question acceptable)												
Human incidence data available?	Yes	1	1	1	1	1	1	1	1	1	1	1
	No			1	1			1	1	1	1	1
Dose-response data available?	Yes	1	1	1	1	1	1	1	1	1	1	1
	No			1	1	1	1	1	1	1	1	1
Occurrence data (concentration, prevalence, dose) available?	Yes	1	1	1	1	1	1	1	1	1	1	1
	No					1	1	1	1	1	1	1
Food consumption data available?	Yes	1	1	1	1	1	1	1	1	1	1	1
	No					1	1	1	1	1	1	1
Growth models available (if no microbiological hazards are ranked, please indicate: not applicable)?	Yes	1	1	1	1	1	1	1	1	1	1	1
	No					1	1	1				
	Not applicable	1	1	1	1	1	1	1	1	1	1	1
Toxicological reference values (ADI, TDI, etc.) available ((if no chemical hazards are ranked, please indicate: not applicable)?	Yes	1	1	1	1	1	1	1	1	1	1	1
	No				1	1	1	1	1	1	1	1
	Not applicable	1	1	1	1	1	1	1	1	1	1	1

3.4.2. Guidelines for the decision tool

In order to determine the most appropriate method for risk ranking, the following steps should be taken:

1. The risk manager should clearly define the aim of the ranking exercise. For example, is the ranking meant to gain scientific insight, allocate scientific budgets or as input to monitoring programs? Furthermore, the hazards to be ranked should be clearly indicated as well as the food products to be included.

2. The risk manager should specify the prerequisites for performing the risk ranking analysis:

   a. how many man hours (or how much budget) are available for data collection and performing the risk ranking analysis?

      i. High (> 12 man months)

      ii. Moderate (6-12 man months)

      iii. Low (< 6 man months)

   b. what type of output is needed?

      i. Quantitative (showing actual numbers)

      ii. Semi-quantitative (based on scores for both severity and probability)

      iii. Qualitative (specifying low, medium, high)

   The amount of resources available and the expected outcome are related. In general, a more quantitative method such as a risk assessment requires a higher amount of resources than a more qualitative method such as a risk matrix.

   c. Is it important that the methodology can be easily explained to stakeholders (laymen)?

   In general, a risk manager should be able to understand the outcome of the risk ranking. However, not all risk ranking methods are evenly easy to understand. Some methods require some in-depth knowledge to fully understand the outcome. If the final ranking of food safety hazards should be easily understandable in order to communicate with stakeholders or the general public, this question should be answered with Yes in order to select the most appropriate method.

   d. should stakeholder perception be included?

   If it is important that stakeholder opinions are included in the analysis, this question should be answered with Yes.

   e. should weights be included for the various risk ranking criteria included in the model?

Potentially each risk ranking method contains several possible parameters and criteria. Stakeholders and experts may have an opinion on the relative importance of including these elements in the risk ranking exercise. If the risk manager thinks it is important to include such weighing in the analysis, the question should be answered with Yes.

f. should human incidences be included?

If human incidences should be included in the analysis or should be part of the output, this question should be answered with Yes.

g. should economic impact be included?

If economic impact should be part of the analysis, the question should be answered with Yes.

h. what is the preferred method of communication?

The outcome of the various risk ranking methods can be graphs, lists, a decision tree or a combination of these. The risk manager should specify the preferred method of output.

After all these questions are answered, they can be filled in into the established Decision Tool in Excel, by including a "1" for all acceptable answers in the yellow column (multiple answers can be indicated per question). The entry code is: EFSA.

3. The risk assessor should try to find as much data as possible in order to rank the specified hazards. The following questions need to be answered by the risk assessor:

a. Are there any human incidence data available for the specified hazards?

If data from monitoring programs or literature studies are available this question should be answered with yes. If these data are not available, it may be possible to estimate them, in case dose response data, occurrence data and food consumption data are available (see following questions).

b. Are dose-response data available for the specified hazards?

If dose-response data are available from literature, for example from animal experiments or epidemiological studies, the question should be answered with Yes.

c. Are occurrence data available for the specified hazards?

If data are available on concentration, prevalence and/or dose of the specified hazards in the food products studied, the question should be answered with Yes

d. Are food consumption data available for the specified hazards?

If food consumption data are available, for example from consumption surveys, the question should be answered with Yes

e. Are growth models available for microbiological hazards?

If growth models are available for the various steps in the food production chain, the question can be answered with Yes. If the risk ranking method is to be applied on non-microbiological hazards, the question should be answered with: Not applicable.

f. Are toxicological data available for the specified hazards?

If ADI, TDI, RfD or other toxicological reference values are available, the question should be answered with Yes. In case the risk ranking method is to be applied on non-chemical hazards, the question should be answered with: Not applicable

Once all these questions are answered, they can be filled in into the established Decision Tool in Excel, by including a "1" for all acceptable answers in the yellow column. The entry code is EFSA. In contrast to the questions on the risk manager's prerequisites, only one answer can be given for each question.

Once all the answers are filled in into the Excel Decision Tool, the evaluation part will show the most appropriate method(s) to perform the risk ranking exercise. It will also indicate why certain methods are excluded: either due to the fact that the prerequisites are not met, or that the data needed to perform the method are not available. In general, methods to the left in the table are time consuming and data demanding, but will result in quantitative outputs, whereas methods on the right can be performed with limited budgets and scarce data sets, but the outcome will be more qualitative. Based on the outcome of the decision tool, the risk manager and the risk assessor can discuss which method to choose for the risk ranking analysis. It is also possible to start filling in the tool with the questions on data availability and see which methods are available based on the available data. The outcome, i.e. the one or more preferred methods, can be used as input for discussion with the risk manager to determine which methods are suitable based on the available data. Combined with the risk manager's prerequisites, the most appropriate method can then be defined.

In order to show how the decision tool can be used, the following section describes the results from a feasibility study, in which several case studies are performed for ranking the risks related to the presence of chemical, microbiological and nutritional hazards. Several assumptions were made regarding the aim and risk manager's prerequisites in order to fill in the decision tool. Therefore, the case studies can only be seen as examples on the use of the decision tool. The inputs used in the decision tool are indicated in Table 4 below. For four case studies, the most appropriate method was also applied to demonstrate how the methods can then be used to actually rank the specified hazards. These case studies were: risk ranking of *Fusarium* mycotoxins in cereals (3.4.3.1), risk ranking of zoonoses in fresh meat (3.4.3.5), risk ranking of nutritional hazards (3.4.3.8) and risk ranking of chemical, microbiological and nutritional hazards using a common denominator (3.4.3.9). The outcome of these risk ranking exercises should merely be seen as illustration on the use of the selected methods in risk ranking the specified hazards.

Table 4: Inputs used for selection of the appropriate methods for use in the various case studies

Question	Answer	Myco-toxins	Pyrro lizidine alkaloids I	Pyrro lizidine alkaloids II	Zoonoses in fresh meat I	Zoonoses in fresh meat II	Viruses in fruits	Botani-cals	Nutritional hazards	Microbio-logical, chemical and nutritional hazards
RISK MANAGER PREREQUISITIVES (Several answers per question acceptable)										
Amount of person hours resources available?	High		1					1	1	1
	Moderate	1		1	1	1	1	1		
	Low					1				
Level of output needed?	Quantitative		1	1			1	1	1	
	Semi-quantitative	1		1	1	1	1	1		1
	Qualitative			1		1		1		
Is it important that the methodology can be easily explained to stakeholders (laymen)?	Yes				1	1	1			1
	No	1	1	1	1	1		1	1	1
Should stakeholder perception be included?	Yes						1	1		
	No	1	1	1	1	1		1	1	1
Should uncertainty be included?	Yes		1		1	1	1	1	1	1
	No	1		1				1		
Should weights be included for the various risk ranking criteria?	Yes						1			
	No	1	1	1	1	1	1	1	1	1
Should human incidences be included?	Yes				1	1	1		1	1
	No	1	1	1		1		1		

Question	Answer	Myco-toxins	Pyrro lizidine alkaloids I	Pyrro lizidine alkaloids II	Zoonoses in fresh meat I	Zoonoses in fresh meat II	Viruses in fruits	Botani-cals	Nutritional hazards	Microbio-logical, chemical and nutritional hazards
Should economic impact be included?	Yes				1		1			
	No	1	1	1	1	1	1	1	1	1
Preferred method of communication	Graph	1	1	1	1	1	1			1
	Tabulation	1	1	1	1	1	1	1	1	1
	Decision Tree	1	1	1	1	1	1	1		1

DATA AVAILABILITY (only one answer per question acceptable)										
Human incidence data available?	Yes				1	1			1	1
	No	1	1	1			1	1		
Dose-response data available?	Yes			1	1	1	1		1	1
	No	1	1				1	1		
Occurrence data (concentration, prevalence, dose) available?	Yes	1		1	1	1	1		1	1
	No		1				1	1		
Food consumption data available?	Yes	1	1	1	1	1	1		1	1
	No							1		
Growth models available ((if no microbiological hazards are ranked, please indicate: not applicable)?	Yes				1	1				1
	No									
	Not applicable	1	1	1			1	1	1	
Toxicological reference values (ADI, TDI, etc.) available ((if no chemical hazards are ranked, please indicate: not applicable)?	Yes	1	1	1						1
	No									
	Not applicable				1	1	1	1	1	

3.4.3. Results of the feasibility study

Various case studies were performed to determine the usefulness of the developed decision tool for choosing the most appropriate risk ranking method. The following paragraphs demonstrate the use of the decision tool for ranking the risks related to the presence of chemical, microbiological and nutritional hazards. In some case studies the selected method was also applied to rank the risks studied: case study on mycotoxins (3.4.2.1), case study on zoonoses (3.4.2.4), case study on nutritional hazards (3.4.2.8) and a case study on ranking the risks of microbiological, chemical and nutritional hazards in food (3.4.2.9).

3.4.3.1.   Case study on mycotoxins

In this case study the most suitable method was selected for risk ranking of the major *Fusarium* mycotoxins. The method was then applied to prioritize the human health risk related to exposure to these mycotoxins via food consumption. This ranking exercise was only performed for illustration purposes.

Aim: to perform a preliminary ranking of the risks to human health related to exposure - via food consumption - to the three mycotoxins of deoxynivalenol, nivalenol and zearalenone, such to efficiently allocate budget for further scientific research.

Risk manager prerequisites: For this case study, the following assumptions were made regarding the risk manager prerequisites:

- time and resources available are moderate,

- no stakeholder input is needed,

- no economic analysis is needed,

- the method should be as accurate as possible,

- therefore important to include uncertainty,

- the method of communication is not important, either graphs, lists or decision trees,

- there is no need to weigh various risk ranking criteria.

Data availability: Literature research showed that there were occurrence data available for the specified mycotoxins in cereal grains (EFSA, 2013; EFSA CONTAM panel, 2011b, 2013). Furthermore, toxicological reference values were available for the selected mycotoxins. Some toxicological information on the effect of mycotoxins were also available.

Outcome decision tool: The interactive Excel decision tool was filled in using the risk manager's prerequisites and data availability as indicated above. This resulted in three possible suitable methods, being risk ratio, scoring method and MCDA (Annex II). As the risk manager indicated that there is no need to weigh the risk ranking criteria, there was no need to perform an MCDA. As the risk ratio method is the least time consuming, this method was chosen to perform a preliminary risk ranking of the specified mycotoxins.

Outcome risk ranking mycotoxins: For this case study, hazard indices were estimated for chronic exposure to deoxynivalenol, zearalenone and nivalenol based on scientific reports from EFSA. Hazard indices were estimated by dividing the median of the estimated daily intakes by the tolerable daily intake. For deoxynivalenol, a provisional maximum tolerable daily intake of 1 µg/kg bw/day was used (EFSA, 2013), for nivalenol and zearalenone the TDI as established by the CONTAM panel was used (1.2 µg/kg bw/day and 0.25 µg/kg bw/day, respectively) (EFSA CONTAM panel, 2011b, 2013). Figure 4 presents the resulting hazard indices, specified per age group. Based on the results of applying this method, the risk related to zearalenone was ranked highest, followed by nivalenol and deoxynivalenol.
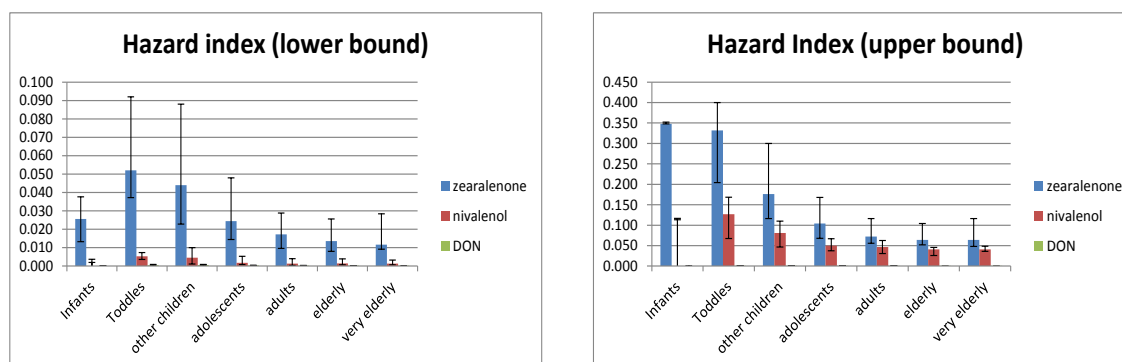
Figure 4: Estimated hazard index based on median estimated daily intakes for several age groups. Error bars reflect the minimum and maximum hazard index. The left figure is obtained by assuming concentrations <LOD are zero and the right figure is obtained by assuming concentrations < LOD are equal to the LOD.

3.4.3.2.   Case study on pyrrolizidine alkaloids, part I

Aim: To perform a ranking of the human health risk related to the most important pyrrolizidine alkaloids (PA) in herbs, honey, milk, eggs and meat in order to efficiently allocate research budget. The following PA (including the tertiary amine as well as the corresponding N-oxide forms) were considered of particular importance for food and feed (EFSA CONTAM panel, 2011a):

- Senecionine-type PA: acetylerucifoline, erucifoline, integerrimine, jacobine, jacoline, jaconine, jacozine, retrorsine, senecionine, seneciphylline,

- Lycopsamine-type PA: acetylechimidine and isomers, echimidine and isomers, echivulgarine, lycopsamine and isomers, vulgarine,

- Heliotrine-type PA: europine, heliotrine, lasiocarpine,

- Monocrotaline-type PA: fulvine, monocrotaline, retusamine, trichodesmine.

Risk manager prerequisites: For this case study, the following assumptions were made regarding the risk manager prerequisites:

- time and resources available are high (> 12 man months),

- no stakeholder input is needed,

- no economic analysis is needed,

- preferably a quantitative output is obtained,

- therefore important to include uncertainty,

- the method of communication is not important, either graphs, lists or decision trees,

- there is no need to weigh various risk ranking criteria,

Data availability: There is some data on the occurrence of PA, primarily in honey. Data is available for retrorsine, senecionine, seneciphylline, acetylchimidine, echimidine,e chivulgarine, lycopsamine, heliotrine and lasiocarpine. As 1,2-unsaturated PA are genotoxic and carcinogenic, no TDIs are available for these PA. However, a BMDL10 has been established of 70 µg/kg bw (EFSA CONTAM panel, 2011a).

Outcome decision tool: Risk assessment is the only method that fulfils the risk manager's requirements (Annex II). As there is hardly any data available on the presence of PAs in various food products and dose response data are limited, it is, currently, not possible to perform a risk assessment on PAs. . Therefore, in order to perform a risk ranking of these chemical compounds, according to the risk manager's prerequisites, more scientific budget should go in obtaining the required data. Another option is to downsize the risk manager's prerequisites in order to find an appropriate method that can be used to rank the risks of PAs, given the data available. In the latter case, if the level of accuracy is downsized to a semi-quantitative outcome, MCDA might be applicable.

3.4.3.3.   Case study on pyrrolizidine alkaloids, part II

In order to show how the outcome of the decision tool may change when the case study is defined with different prerequisites, the following case study has been drawn up.

Aim: To perform a ranking of the human health risk related to PA in honey for which occurrence data are available to efficiently allocate research budget. The following PA were incorporated in this study: retrorsine, senecionine, seneciphylline, acetylchimidine, echimidine,echivulgarine, lycopsamine, heliotrine and lasiocarpine.

Risk manager prerequisites: For this case study, the following assumptions were made regarding the risk manager prerequisites:

- time and resources available are moderate (6- 12 man months),

- no stakeholder input is needed,

- no economic analysis is needed,

- no preference for the obtained accuracy,

- no need to include uncertainty,

- the method of communication is not important, either graphs, lists or decision trees,

- there is no need to weigh various risk ranking criteria,

Data availability: Occurrence data is available for the specified PA. Furthermore, a $BMDL_{10}$ has been established of 70 µg/kg bw (EFSA CONTAM panel, 2011a).

Outcome decision tool: The following methods fulfilled the risk manager's requirements: ratio, scoring, MCDA, risk matrix, flow charts/decision trees and expert judgment (Annex II). The first three methods result in more accurate outcomes than the latter three. As the risk manager prerequisites indicate there is no need to weigh various risk ranking criteria, there is also no need to perform an MCDA. Either risk ratio or scoring methods can then be used to perform the risk ranking study, in which the first method is less time consuming. The EFSA CONTAM panel decided to perform a ratio method by estimating the Margin of Exposure (MoE) for the specified PA in honey for various age groups. For example, the MoE for adults ranged between 57,000-3,500,000 and between 7400->7,000,000 at the mean and 95th percentile of consumption (based on maximum UB and minimum LB across European countries) (EFSA CONTAM panel, 2011a).

3.4.3.4. Case study botanicals

Aim: Although efforts are being made to systematically acquire and organise information about botanicals (e.g., in the EU project 245199, PlantLIBRA), for many botanicals there is a lack of necessary information for risk assessment. This is demonstrated in the 6 case studies examined by ESCO (2009).

To perform a ranking of botanicals of concern in order to prioritize European research funding for safety assessments of botanicals. According to the Compendium from EFSA (2012a), there are known to be around 900 plants which may be of concern. The number of botanical substances of concern could be much larger as different parts of the same plant (roots, leaves etc.) may exhibit different properties, and different production processes may produce different biochemical properties.

Risk manager prerequisites: For this case study, the following assumptions were made regarding the risk manager prerequisites:

- time and resources available are moderate to high (6 months or higher),

- stakeholder input may be needed,

- economic analysis is not needed,

- the output may be quantitative or qualitative,

- estimates of uncertainty can be included,

- method of communication is not important.

Data availability: As a consequence of the data limitations, it is not possible to rank the risk to human health (or net benefits) on the basis of their socio-economic impacts. There is a lack of epidemiological data. Furthermore, there is a lack of occurrence data and toxicological information is not available for all botanicals.

Outcome decision tool: When this information is entered into decision tool for choice of available risk ranking methodology, the MCDA, risk matrix, flow charts/decision trees and expert judgment are the most promising methods for risk ranking of botanicals based on the prerequisites and data availability (Annex II). Depending on the preferred outcome, the risk manager may choose a risk matrix when favouring visual representation, or flow charts/decision trees if this is the preferred outcome. MCDA or expert judgment appear promising as they allow weighing up the relative advantages and disadvantages of consuming botanicals. Choice experiments could be held, with its design including attributes upon which participants base their choice decisions: reductions in illness due to ingestion of the individual botanical, cost of purchasing the botanical, and increase in illness due to side-effects of ingestion. However, the required data are believed to be missing. Furthermore, this method is very labour intensive and thus, does not fulfil the risk manager's requirements.

Multi-criteria decision analysis (MCDA) is a flexible tool which has the advantage of being able to integrate different data types, including technical, economic and public perception data, and both quantitative and semi-quantitative data. Broadly, MCDA assigns and then synthesises various criterion scores (either weighted or unweighted) for each object being assessed, which are then ranked on the basis of their overall scores. For botanicals the lack of adequate and consistent data for each substance to be ranked appears to preclude this approach.

Judgement-based approaches based on the current state of knowledge and expert opinion offer a way of overcoming these disadvantages and producing a ranking. This ranking would not be static but would change as new data become available. As well as being used to rank risks or net benefits, it could help identify priorities for further data collection. The proposed method is based on expert deliberation in a workshop format preceded by an expert workshop to define the risks which are to be considered. Information on how to design an expert workshop to prioritize botanicals can be found in Annex III.

3.4.3.5.   Case study zoonoses in fresh meat, part I

Aim: to perform a ranking of bacterial zoonoses and zoonotic agents monitored in fresh pork, beef and broiler meat in Member State (MS) X, in order to evaluate if state and industry control efforts and financial resources are allocated optimally. Zoonoses to be included are food related disease where national surveillance data may be reported to EFSA for the EU Summary Report.

Risk manager prerequisites:

- time and resources available are moderate (6-12 man-months),

- the method should be as accurate as possible, considering data and resources (moderate-high),

- thus estimates of the uncertainty must be included,

- complexity of method is not important. Preferably an easily understandable method is used, but this is not a prerequisite,

- industry stakeholder perception is not needed,

- estimates of incidences of human illness is needed,

- economic impact could also be relevant but not needed,

- weighting should not be applied,

- method of communication of results not important.

Data availability: EU Member States reports national findings of zoonoses and zoonotic agents in food and human domestic outbreaks to EFSAs Zoonosis Monitoring Database. Consumption data is collected by EFSA in the European Food Consumption Database. Human cases are reported to the ECDC database Tessy. Literature can provide additional data and dose-response relationships if required.

Outcome decision tool: When these information is entered into the excel decision tool, DALY/QALY, Cost of Illness methods and Multi Criteria Decision Analysis are the risk ranking methods listed as possibly appropriate (Annex II). As there is not sufficient resources allocated for including Risk Assessments, modelling of incidences of disease for estimating DALYs and Cost-of-Illness could preferably be made in the ECDC Burden of Disease tool BCoDE. In this tool, assumptions about under-reporting and probabilities of health outcomes are very general, some loss of accuracy must be expected compared to constricting a Burden of Disease model based on national data.

If lower accuracy and less easy understandable output are acceptable, MCDA could be an option. Here, data on human incidences could be included in the format also reported to the ECDC database Tessy (reported cases, hospitalisation and/or deaths). Different parameters describing economic impact could also be included if available. The Risk Assessment and Comparative Risk Assessment methodologies were excluded due to the requirement of high resources, whereas methods using Risk Matrix, Flow charts /Decision trees and Expert judgment were excluded due to relatively low accuracy and no representation of uncertainty. Methods such as estimating Ratios, Scoring as well as the Risk Matrix do not include human incidences in the modelling.

Application of the most promising methods (DALY/QALY and Cost of Illness) on the ranking of foodborne diseases in the Netherlands in 2011

Since 2008, The Dutch National Institute for Public Health and the Environment (RIVM) regularly publishes estimates of the incidence, burden and costs of 14 enteric pathogens (Bouwknegt et al., 2013; Mangen et al., 2013). For each enteric pathogen, the overall disease burden (in DALYs) and Cost-of-Illness (in mill. €) was estimated, and then attributed to five major pathways (food, environment, direct animal contact, human–human transmission, and travel) and 11 groups within the food pathway.

Attribution of human cases (and thus the related disease burden) was based on estimates from an expert elicitation (Havelaar et al., 2008). Among the incidences from 11 of the pathogens were attributed to food (*Campylobacter* spp, *Listeria monocytogenes*, Pathogenic *E. Coli* - STEC O157,

Salmonella spp, *Cryptosporidium parvum*, *Giardia Iambiia*, Norovirus as well as toxins from *Bacillus cereus*, *Clostridium perfringens* and *Staphylococcus aureus*). In the reports, results were presented in tables presenting estimated total incidence, deaths, disease burden (DALYs), Cost-of-Illness (Mill. €) for each relevant combination of pathogen and food pathway.

Figure 5 presents a ranking of the 15 combinations of pathogen and meat type with the highest estimated burden of disease and cost of illness. The results indicate the mean values of the estimated DALYs and CoI. Uncertainty analysis has been performed in the assessment, but these results were not available to include in Figure 5. When ranking is based on DALYs, risks related to *Toxoplasma* in pork and *Campylobacter* in poultry clearly rank highest. When ranking is based on the costs related to these food related diseases, *Campylobacter* in poultry is ranked first followed by *Clostridium* toxins in beef and lamb meat, whereas *Toxoplasma* in pork follows as third in rank.

Disability Adjusted Life Year (DALYs)
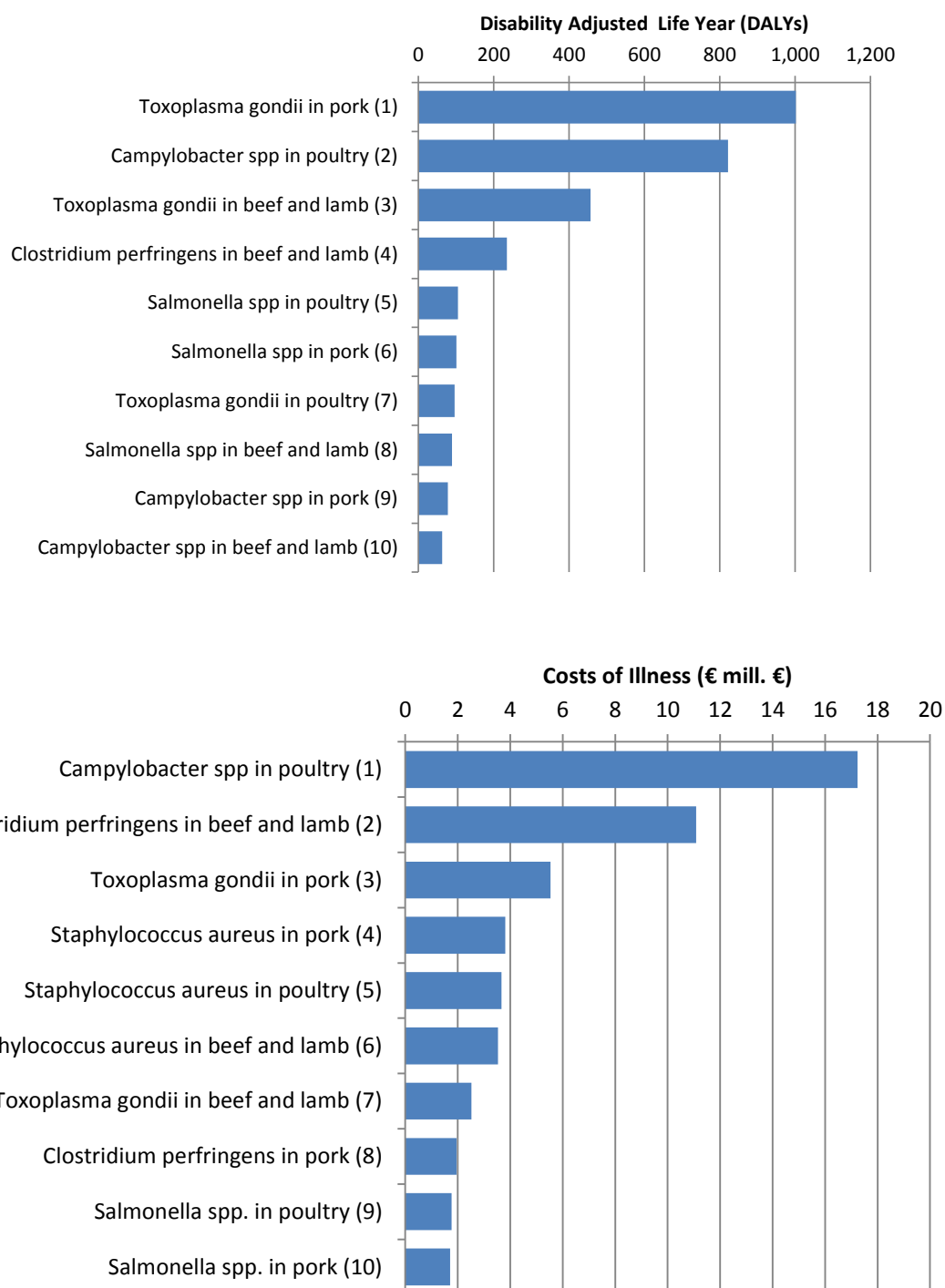

Costs of Illness (€ mill. €)

Figure 5: Ranking of the 15 combinations of pathogen and meat type with the highest estimated attributable disease burden (DALY per year, undiscounted) and costs of illness (Mill. €, discounted) in the Netherlands 2011, based on data from Bouwknegt et al. (2013) and Mangen et al. (2013).

Based on the decision tool an applicable method can be derived. However, risk ranking based on the results as for those from Dutch studies can only be done within the allocated resources, if a general capacity for calculating burden of disease and cost-of-illness is already well established. Alternatively, calculations of DALYs could be made in the ECDC Burden of Disease tool BCoDE. If the Dutch source attribution approach is to be applied in other MSs, the applicability of attribution estimates must be evaluated in a national context.

3.4.3.6. Case study zoonoses in fresh meat, part II

Aim: To perform a risk ranking of bacterial zoonoses and zoonotic agents monitored in fresh pork, beef and broiler meat in Member State X, to help prioritise national research funding.

Risk manager prerequisites: For this case study, the following assumptions were made regarding the risk manager prerequisites:

- time and resources available are low to moderate (3-12 man-months),

- qualitative or semi-quantitative outputs acceptable, but estimates of uncertainty must be included,

- weighting may be applied,

- industry stakeholder input regarding is not needed,

- human incidences may be included (not required), but information reported to ECDC is sufficient,

- economic analysis is not needed,

- method of communication is not important.

Data availability: As in case study zoonoses in fresh meat, part I.

Outcome decision tool: Using a different set of requirements as in part I, shows that several methods may be appropriate for risk ranking: ratio, scoring, DALY/QALY, MCDA and expert judgment (Annex II). As indicated earlier, the ratio and scoring methods do not include human incidences. Therefore, in case human incidences as reported to ECDC are to be included, DALY/QALY, MCDA and expert judgment may be used.

If the request regarding estimates of uncertainty may be disregarded, Flow charts or Decision tree methods may also be an option. The EFSA BIOHAZ panel primarily used this approach to rank the public health hazards to be covered by inspection of meat from poultry (EFSA, 2012d) and pork (EFSA, 2012e).

3.4.3.7.   Case study foodborne viruses in berry fruits and leafy green vegetables

Aim: Contamination of berry fruits and leafy green vegetables by viruses is dominated by a few species, in particular: Norovirus (in particular genogroups GI and GII) and Hepatitis A3 (EFSA, 2011), whereas Hepatitis E appears to be an emerging risk for berry fruits and leafy green vegetables (Kokkinos et al., 2012; Maunula et al., 2013). The aim of this case was to identify specific priority virus-commodity combinations, including consideration of the target population in order to respond to relevant question(s) from the European Commission.

Risk manager prerequisites: The following assumptions were made regarding the risk manager prerequisites:

- Moderate resources are available – therefore we assumed that this was a long-term policy making exercise (not an emergency incident),

- Level of accuracy of the assessment is set at high or moderate meaning that quantitative or semi-quantitative output is required,

- Stakeholders' perceptions are important for the assessment/evaluation.

Data availability: After reviewing the available information, from a short literature review and discussions with experts, the following inputs were obtained:

- Human incidence data is not available.

- Dose response models are available for HAV and Norovirus, but unlikely for HEV.

- Intake (i.e. consumption) data is available for some EU Member States, and extrapolation is assumed for the ones for which data is not available.

- Questions related to growth models and toxicological reference values question are not applicable to viruses.

Outcome decision tool: The particular case study is characterised by (a) lack of crucial data, and (b) variability in the spectrum of missing information amongst the different foodborne viruses. For example, epidemiological data for Norovirus outnumber the available data on HAV and HEV. The variability in the abundance of reliable relevant data inevitably affects the decision on which of the remaining risk ranking methods to employ. The lack of epidemiological data, source attribution data, occurrence data and virus behaviour data (e.g. transfer rate, persistence, shedding) hinders the application of the cost of illness, WTP and DALY/QALY methods. Because of this reason, the even

---

[3] All HAV virus strains are classified into one of six genotypes. Genotypes I, II and III have been isolated from humans, and genotypes IV, V and VI are of simian origin. Genotypes IV, V and VI are genetically distinct from human HAV strains. Genotype I can be divided into two sub-genotypes IA and IB. This genotype is the most prevalent worldwide.

more data-hungry methods, in particular the methods based on exposure assessments (i.e., ratio and scoring methods), and risk assessments, cannot be applied either.

Given the risk manager's prerequisites and the data availability, the decision tool indicates that MCDA is the most appropriate method for risk ranking of the specified foodborne viruses (Annex II). This method category was deemed as appropriate to rank different combinations of foodborne viruses/berry fruits and leafy green vegetables, as it would allow the entries of diverse information, and could tolerate information inputs at a higher (i.e., less detailed) level. Depending on the particular MCDA approach that one chooses to follow the criteria included, and the type (i.e., qualitative, quantitative) of inputs needed would change. An important aspect of any method of MCDA is that expert judgment is always involved, albeit in different formats again depending on the particular approach followed. Therefore, one would need to anticipate the required resources, i.e., time and cost, to execute the necessary elicitation of expert knowledge protocols. A crucial difference amongst different approaches of MCDA concern the way that criteria are weighted. Most often, criteria are either assumed as equal (therefore all given the same weight) or weights are introduced by the users of the approach or experts employed in a MCDA exercise. The former allows for an easier aggregation of different expert views, whereas the latter allows for acquiring information on which criteria are deemed as most important for each expert. In Annex IV, more details on two possible MCDA approaches that reflect the main differences amongst the different MCDA methodologies mentioned above are provided. These include one more traditional MCDA that is based on scoring (EFSA, 2013) and one that is based on probabilistic inversion modelling (Flari et al., 2011; Havelaar et al., 2010; Neslo and Cooke, 2011). The latter is much more labour intensive, but it would carry more value in terms of employing quantitative criteria and avoiding biases, therefore enhancing its transparency, reproducibility, and predictability.

From this case study, it was concluded that there is a distinct lack of data available in order to produce an accurate and reliable risk ranking of foodborne viruses in berry fruits and leafy green vegetables. However, as more and more analytical data become available either from outbreaks or research studies, including surveys of foodstuffs, the more accurate and reliable any risk ranking model will become. The latter is much more applicable to the MCDA approach based on probabilistic inversion. A full risk ranking exercise will be needed to evaluate the (a) applicability of the MCDA based on probabilistic inversion approach in such problems, i.e., foodborne viruses in berry fruits and leafy green vegetables, and (b) the appetite of risk managers for such approach.

3.4.3.8.   Case study nutritional hazards

Due to the very few studies of risk ranking of nutritional hazards and the extensive amount of epidemiological data required for this task, the study described by Danaei et al. (2009) was used as demarcation case study for risk ranking of nutritional hazards. These authors aimed to estimate the mortality effects of a total of 12 modifiable dietary, lifestyle and metabolic risk factors in the US using consistent and comparable methods. This demarcation study will use the data available for Danaei et al. (2009) on dietary risk factors and the decision tool to select the best suited method for risk ranking the selected risk factors.

Aim: To perform a ranking of the following dietary risk factors: high dietary trans fatty acids; low dietary poly-unsaturated fatty acids (PUFA); low dietary omega-3 fatty acids (seafood); high dietary salt (sodium) and low intake of fruit and vegetables based on mortality data by age and sex in order to support priority setting and policy making within the US.

Risk manager prerequisites: For this case study, the following assumptions were made regarding the risk manager prerequisites:

- time and resources available are high,

- no stakeholder input is needed,

- no economic analysis is needed,

- the method should be as accurate as possible,

- therefore important to include uncertainty,

- the method of communication should preferable be presented as a prioritized list.

Data availability: Relevant data is available from the following sources:

- data on population exposures and disease-specific deaths from national surveys are available,

- well conducted epidemiological studies are available from the published literature,

- Intake data: risk factor exposure data in the US population from nationally representative health surveys,

- Human incidence data: disease-specific mortality statistics from national health statistics,

- Dose-response data: etiological effects of risk factors on disease-specific mortality from systematic reviews and meta-analysis of epidemiological studies.

Outcome decision tool: When the available information is entered into the decision tool for selection of available risk ranking methods, three methods being risk assessment, comparative risk assessment (CRA) analysis and DALY/QALY are listed as applicable (Annex II). All these methods require high resources, offer high accuracy, includes uncertainty and can be presented in the preferred format.

According to our aim a ranking should be performed on mortality from the risk factors and, therefore, a CRA analysis, which estimates the number of disease-specific deaths attributable to all non-optimal levels of each risk factor exposure, was chosen. The result from the CRA analysis is a list of deaths from all causes (thousands of deaths) attributable to the specific dietary risk factors and the 95% confidence intervals. An example with total deaths (95% confidence intervals) for females from the five dietary risk factors is presented in Table 5 below, adapted from Danaei et al. (2009).

Example output from the comparative risk analysis for total deaths (with 95% confidence intervals) for females from each of five dietary risk factors, adapted from Danaei et al. (2009).

Risk factor	Females total
High dietary salt (sodium)	54 (50-57)
Low dietary omega-3 fatty acids (seafood)	39 (31-47)
High dietary trans fatty acids	35 (23-46)
Low intake of fruits and vegetables	24 (15-36)
Low dietary PUFA	6 (3-9)

3.4.3.9.   Case study on ranking microbiological, chemical and nutritional risks in food (based on common denominator)

In this study the ranking of a broad range of risks (due to different types of hazards) was studied. Based on the risk manager's prerequisites and the available data, the most appropriate method was determined for human health risks related to food intake.

Aim: to compare the human health risks related to chemical, microbiological and nutritional hazards in food in order to prioritise governmental policy.

Risk manager prerequisites: For this case study, the following assumptions were made regarding the risk manager prerequisites:

-   time and resources available is high (> 12 man months),

-   the method should be as accurate as possible, considering data and resources ((semi-) quantitative),

-   thus estimates of the uncertainty must be included,

-   complexity of method is not important. Preferably an easily understandable method is used, but this is not a prerequisite,

-   industry stakeholder perception is not needed,

-   estimates of incidences of human illness is needed,

-   economic impact is not needed,

-   weighting should not be applied,

- method of communication of results not important.

Data availability: for a number of microbiological, chemical and nutritional hazards in food data is available on occurrence, human incidence and dose response. Furthermore, toxicological reference values are available for most chemical hazards.

Outcome decision tool: When the risk manager's prerequisites and the availability of data is filled in into the decision tool, both MCDA and DALY/QALY appear to be the most appropriate methods for comparing various risks based on a common denominator (Annex II). DALY/QALY allows for comparing the risks related to various types of food hazards as well as relating these risks to other factors that influence human health (such as smoking or traffic incidents).

As DALY/QALY seemed to be the most appropriate method, we searched for an example in which this method was applied for ranking microbiological, chemical and nutritional hazards. The RIVM (the National Institute of Public Health and the Environment in the Netherlands) has estimated DALYs for various microbiological, chemical and nutritional risks related to food intake in the Netherlands (van Kreijl et al., 2004). Table 6 gives a risk ranking of nutritional, microbiological and chemical hazards based on DALYs.

The table shows that the health loss due to an unfavourable nutritional composition and by obesity is about two orders of magnitude greater than the health loss due to microbiological contamination. Comparison between the risks of microbiological and chemical hazards shows that the latter causes lower DALYs. Governmental policy should thus aim at improving a healthy diet focusing on the quality as well as the quantity of food intake (in order to prevent obesity) (van Kreijl et al., 2004). Although nutritional risks are ranked as most important, the number of relevant references describing a ranking method for nutritional related risk is, currently, limited. Therefore, more research is needed in prioritising human health risks related to nutritional hazards.

Table 5: Ranking of risks related to nutritional, microbiological and chemical hazards in food based on estimated DALYs (van Kreijl et al., 2004).

Number of lost DALYs	Nutritional hazards	Microbiological hazards	Chemical hazards
>300,000	Too many SFA[1], TFA[2], too few fruits, vegetables and fish and a positive energy balance (causing obesity)		
100,000-300,000	Too many SFA[1], TFA[2], too few fruits, vegetables and fish		
30,000-100,000	Too many trans fatty acids, too few fruits, vegetables and fish		
10,000-30,000	Too many saturated fatty acids		
3,000-10,000		Gastroenteritis by microorganisms in food	
1,000-3,000			
300-1,000		*Campylobacter* in food	Allergens, acrylamide
<300		STEC O157[3]	PAHs[4], other studied compounds (mycotoxins, phycotoxins, phytotoxins, clenbuterol and nitrate)

[1]SFA: saturated fatty acids; [2] TFA: trans fatty acids; [4]STEC: shiga toxin producing *Escherichia coli*; [3]PAHs: polycyclic aromatic hydrocarbons

This case study showed that the DALY/QALY concept can be used as a method for comparing the risks of various types of hazards, although it is not always possible to derive DALY/QALYs for all food safety risks. Especially for chemical hazards it is difficult to relate the intake of chemical compounds to human health loss as many compounds have a chronic rather than an acute effect. For example, the report showed that it was not possible to derive a DALY for PCBs and dioxins. Using

some assumptions related to human incidences and weighing factors, it was possible to derive DALYs for the other studied chemical hazards in the Netherlands as listed in table 6 (van Kreijl et al., 2004). For further details on the establishment of the DALYs as well as uncertainties encountered, we refer to the report of the RIVM (van Kreijl et al., 2004).

In the Netherlands, relevant data were available in order to derive DALYs for the various risks. For other MSs, an approach could be to calculate the DALYs to the population in question by working out estimates for number of cases with disease due to each risk factor and then multiply the incidences with the published estimates of DALYs per case (average or distribution).

The DALY approach has also been used by Newsome (2009) and has been incorporated in a web-based approach by the FDA (iRisk) (Chen et al., 2013) in order to prioritize the risks related to the presence of chemical and microbiological hazards in food. The method has also been applied to rank the risks of a range of chemical residues and contaminants (Crawford-Brown, 2012).

4.   Conclusions and recommendations

An extensive systematic review has been performed on methods for ranking of risks related to feed/food safety and nutritional hazards, on the basis of human health impact. Results showed that a wide range of studies apply risk ranking methods using a similar approach but using specific inputs for the purpose of the study. Comparing the relevant references showed that methods can be grouped into some main categories. All method categories have their own specifications for use. Based on the grouping, an overarching framework tool has been developed, which may help the risk assessor, together with the risk manager, to choose the most suitable method. The methods range from very extensive in terms of resources (data, man hours etc.) needed - in case quantitative outcomes are needed - to methods that can be used with time and data constraints, which will result in more qualitative outcomes. Depending on the preferred outcome, a different risk ranking method can be selected for ranking hazards within one hazard category (either chemical, microbiological or nutritional hazards). Risk assessments and CRA are the best methods for providing quantitative results, considering uncertainty and variability in the input data. More qualitative methods could be used when data are scarce for example when emerging risks, such as botanicals, are to be ranked. In these cases of limited data availability, the appropriate methods are MCDA, risk matrix, flow charts/decision trees and expert judgment with an emphasis on input from experts. It is not possible to pinpoint one method that is most suitable for ranking hazards within a hazard category. Choosing the most appropriate method depends on various factors. These factors formed the basis of the developed decision framework. For example, when the risk managers want to base the risk ranking on the basic definition of risk, in terms of prevalence times severity, then the following methods can be applied: RA, CRA, ratio method, scoring method, DALY/QALY, MCDA, risk matrix, flow charts and expert judgment. When it is important to take into account other factors, such as human perception and/or economic impact, in addition to human health impact, then MCDA seems the most appropriate method as it can include a broad range of factors and allows for weighing their importance. WTP can be used when only perception of the risks is considered, and CoI in case only economics are considered.

When hazards are to be ranked between hazard categories (comparing nutritional, chemical and/or microbiological) based on a common denominator, the DALY/QALY concept and Cost of Illness

seem to be most appropriate. However, it should be mentioned that it may not be possible to derive DALY/QALY for all food safety hazards of interest. As indicated before, estimating DALYs for chemical hazards is more difficult than for microbiological hazards as there is often a less clear relationship between intake of the hazard and human health consequences. There is thus more uncertainty in the establishment of DALYs for chemical hazards, which should be taken into account when comparing food safety hazards between hazard categories. Willingness to pay and expert judgment can also be used to rank nutritional, chemical and microbiological hazards, in particular when data availability is low, but these methods are less objective.

Results from the feasibility study showed that the interactive decision tool is useful for selection of the most appropriate methods for the risk ranking task and risk manager's pre-requisites at hand. Usually, it will provide more than one suitable method. There is not one specific risk ranking method that is the best method for all cases. Each of the methods has its own strengths and weaknesses, which are described in this report. The methods should be further explored to determine whether all necessary data are available for the hazards to be ranked in order to use the particular method. Also, insight into the risk manager's preferences in order to select the preferred method is necessary. Therefore, a close cooperation between risk manager and risk assessor is needed to come to the most suitable method for risk ranking.

The project results and developed decision tool are expected to be valuable for EFSA and other risk managers/assessors for selecting the most appropriate methods for risk ranking of feed/food and diet related hazards, on the basis of human health impact. Usually, the most appropriate method for risk ranking is chosen based on expertise and preferences within the working group performing the risk ranking exercise. The developed decision tool will facilitate this decision process and allows for a structured and transparent selection of the most appropriate risk ranking method. To our knowledge, this is the first tool available for this purpose.

REFERENCES

Aenishaenslin C, Hongoh V, Cisse HD, Hoen AG, Samoura K, Michel P, Waaub JP and Belanger D, 2013. Multi-criteria decision analysis as an innovative approach to managing zoonoses: results from a study on Lyme disease in Canada. BMC Public Health, 13.

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

Arnot JA and Mackay D, 2008. Policies for chemical hazard and risk priority setting: Can persistence, bioaccumulation, toxicity, and quantity information be combined? Environmental Science and Technology, 42, 4648-4654.

Arnot JA, Mackay D, Webster E and Southwood JM, 2006. Screening level risk assessment model for chemical fate and effects in the environment. Environmental Science & Technology, 40, 2316-2323.

Aschberger K, Micheletti C, Sokull-Klüttgen B and Christensen FM, 2011. Analysis of currently available data for characterising the risk of engineered nanomaterials to the environment and human health - Lessons learned from four case studies. Environment International, 37, 1143-1156.

Aylward LL, Kirman CR, Schoeny R, Portier CJ and Hays SM, 2013. Evaluation of biomonitoring data from the CDC national exposure report in a risk assessment context: Perspectives across chemicals. Environmental Health Perspectives, 121, 287-294.

Bang DY, Kyung M, Kim MJ, Jung BY, Cho MC, Choi SM, Kim YW, Lim SK, Lim DS, Won AJ, Kwack SJ, Lee Y, Kim HS and Lee BM, 2012. Human Risk Assessment of Endocrine-Disrupting Chemicals Derived from Plastic Food Containers. Comprehensive Reviews in Food Science and Food Safety, 11, 453-470.

Baptista FM, Alban L, Olsen AM, Petersen JV and Toft N, 2012. Evaluation of the antibacterial residue surveillance programme in Danish pigs using Bayesian methods. Preventive Veterinary Medicine, 106, 308-314.

Barendsz AW, Van der Linden AFC and Evers AJ, 1997. Fist-Risk: een softwaremodule voor risicotaxatie: HACCP en risicotaxatie. Voedingsmiddelentechnologie, 30, 16-18.

Batz MB, Hoffman S and Morris Jr JG, 2011. Ranking the Risks: The 10 Pathogen-Food Combinations With The Greatest Burden on Public Health. In (pp. 70 p.): University of Florida.

Batz MB, Hoffmann S and Morris Jr JG, 2012. Ranking the disease burden of 14 pathogens in food sources in the united states using attribution data from outbreak investigations and expert elicitation. Journal of Food Protection, 75, 1278-1291.

Bietlot HP and Kolakowski B, 2012. Risk assessment and risk management at the Canadian Food Inspection Agency (CFIA): A perspective on the monitoring of foods for chemical residues. Drug Testing and Analysis, 4, 50-58.

Bogaardt M-J, Mangen M-JJ, De Wit GA, Nauta MJ and Havelaar AH, 2004. Controlling Campylobacter in the chicken meat chain - Towards a decision support model. In (pp. 56 p.). Bilthoven, The Netherlands: RIVM.

Bonano EJ, Apostolakis GE, Salter PF, Ghassemi A and Jennings S, 2000. Application of risk assessment and decision analysis to the evaluation, ranking and selection of environmental remediation alternatives. Journal of Hazardous Materials, 71, 35-57.

Bouwknegt M, Friesema IHM, Van Pelt W and Havelaar AH, 2013. Disese burden of food-related pathogens in the Netherlands, 2011. In RIVM letter report 330331006/2013.

Bu Q, Wang D and Wang Z, 2013. Review of screening systems for prioritizing chemical substances. Critical Reviews in Environmental Science and Technology, 43, 1011-1041.

Buzby JC, Roberts T, Lin CTJ and MacDonald JM, 1996. Bacterial Foodborne Disease: Medical Costs and Productivity Losses. In: US Department of Agriculture.

Calliera M, Finizio A, Azimonti G, Benfenati E and Trevisan M, 2006. Harmonised pesticide risk trend indicator for food (HAPERITIF): The methodological approach. Pest Management Science, 62, 1168-1176.

Cardoen S, Van Huffel X, Berkvens D, Quoilin S, Ducoffre G, Saegerman C, Speybroeck N, Imberechts H, Herman L, Ducatelle R and Dierick K, 2009. Evidence-based semiquantitative methodology for prioritization of foodborne zoonoses. Foodborne Pathogens and Disease, 6, 1083-1096.

Charles KJ and Ashbolt NJ, 2009. Quantitative microbial risk assessment: a catchment management tool to delineate setback distances for septic systems. (Water and Environmental Management Series). In. London: IWA Publishing.

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.

Chilton S, Covey J, Hopkins L, Jones-Lee M, Loomes G, Pidgeon N and Spencer A, 2002. Public Perceptions of Risk and Preference-Based Values of Safety. 25, 211-232.

Clarke BO and Smith SR, 2011. Review of 'emerging' organic contaminants in biosolids and assessment of international research priorities for the agricultural use of biosolids. Environment International, 37, 226-247.

Codex Alimentarius, 2012. Principles and guidelines for the conduct of microbiological risk assessment. CAC/GL 30-1999 (modified in 2012). In.

Cope S, Frewer LJ, Renn O and Dreyer M, 2010. Potential methods and approaches to assess social impacts associated with food safety issues. Food Control, 21, 1629-1637.

Crawford-Brown DC-B, S., 2012. Cumulative risk assessment framework for waterborne contaminants. Journal of Environmental Protection, 3, 400-413.

Crettaz P, Pennington D, Rhomberg L, Brand K and Jolliet O, 2002. Assessing human health response in life cycle assessment using ED10s and DALYs: Part 1 - Cancer effects. Risk Analysis, 22, 931-946.

Critto A, Torresan S, Semenzin E, Giove S, Mesman M, Schouten AJ, Rutgers M and Marcomini A, 2007. Development of a site-specific ecological risk assessment for contaminated sites: Part I. A multi-criteria based system for the selection of ecotoxicological tests and ecological observations. Science of the Total Environment, 379, 16-33.

Crutchfield SR, Buzby JC, Roberts T and Ollinger M, 1999. Assessing the costs and benefits of pathogen reduction. Food Safety, 22, 6-9.

Dabrowski JM, Shadung JM and Wepener V, 2014. Prioritizing agricultural pesticides used in South Africa based on their environmental mobility and potential human health effects. Environment International, 62, 31-40.

Danaei G, Ding EL, Mozaffarian D, Taylor B, Rehm J, Murray CJL and Ezzati M, 2009. The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors. PLoS Medicine, 6.

DeKay ML, Fishbeck PS, Florig HK, Morgan MG, Morgan KM, Fischhoff B, Jenni KE, Hoffmann S and Taylor MR, 2005. Judgement-based risk ranking for food safety. In S. Hoffmann & M. R. Taylor (Eds.), Toward safer food. Perspectives on risk and priority setting. (pp. 198-226). Washington DC, USA: Resources for the future.

Dreyer M, Renn O, Cope S and Frewer LJ, 2010. Including social impact assessment in food safety governance. Food Control, 21, 1620-1628.

Dybing E, O'Brien J, Renwick AG and Sanner T, 2008. Risk assessment of dietary exposures to compounds that are genotoxic and carcinogenic-An overview. Toxicology Letters, 180, 110-117.

EFSA (European Food Safety Authority), 2010. Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal 2010, 8, 1637.

EFSA (European Food Safety Authority), 2011a. Guidance for risk assessment of food and feed from genetically modified plants. EFSA Journal 2011, 8, 2150.

EFSA (European Food Safety Authority), 2011b. Overview of the procedures currently used at EFSA for the assessment of dietary exposure to different chemical substances. EFSA Journal 2011, 9. 2490

EFSA (European Food Safety Authority), 2012a. Compendium of botanicals reported to contain naturally occurring substances of possible concern for human health when used in food and food supplements. EFSA Journal 2012, 10, 2663.

EFSA Panel on Biological Hazards (BIOHAZ), 2012b. Scientific Opinion on Reflecting on the experiences and lessons learnt from modelling on biological hazards. EFSA Journal 2012, 10, 2725.

EFSA Panel on Biological Hazards (BIOHAZ), 2012c. Scientific Opinion on the development of a risk ranking framework on biological hazards. EFSA Journal 2012, 10, 2724.

EFSA Panel on Biological Hazards (BIOHAZ), 2012d. Scientific Opinion on the public health hazards to be covered by inspection of meat (poultry). EFSA Journal 2012, 20, 2741.

EFSA Panel on Biological Hazards (BIOHAZ), 2012e. Scientific Opinion on the public health hazards to be covered by inspection of meat (swine). EFSA Journal 2012, 20, 2351.

EFSA (European Food Safety Authority), 2013. Deoxinivalenol in food and feed: occurrence and exposure. EFSA Journal 2013, 11, 3379.

EFSA Panel on Contaminants in the Food Chain (CONTAM), 2011a. Scientific opinion on pyrrolizidine alkaloids in food and feed. EFSA Journal 2011, 9, 2406.

EFSA Panel on Contaminants in the Food Chain (CONTAM), 2011b. Scientific opinion on the risks for public health related to the presence of zearalenone in food. EFSA Journal 2011, 9, 2197.

EFSA Panel on Contaminants in the Food Chain (CONTAM), 2013. Scientific opinion on risks for animal and public health related to the presence of nivalenol in food and feed. EFSA Journal 2013, 11, 3262.

Eisenberg JNS and McKone TE, 1998. Decision tree method for the classification of chemical pollutants: Incorporation of across-chemical variability and within-chemical uncertainty. Environmental Science and Technology, 32, 3396-3404.

Ellis B and Garelick H, 2008. A multi-criteria approach for assessing options to remediate arsenic in drinking water. Reviews of Environmental Contamination & Toxicology, 197, 129-161.

Emmanuel E, Pierre MG and Perrodin Y, 2009. Groundwater contamination by microbiological and chemical substances released from hospital wastewater: Health risk assessment for drinking water consumers. Environment International, 35, 718-726.

ESCO, 2009. Advice on the EFSA guidance document for the safety assessment of botanicals and botanical preparations intended for use as food supplements, based on real case studies. EFSA Journal, 7, 280-.

Etchie AT, Etchie TO, Adewuyi GO, Krishnamurthi K, Devi SS and Wate SR, 2013. Prioritizing hazardous pollutants in two Nigerian water supply schemes: A risk-based approach. Bulletin of the World Health Organization, 91, 553-561J.

Evers EG and Chardon JG, 2010. A swift quantitative microbiological risk assessment tool. Food Control, 21, 319-330.

FAO, 1998. Food Quality and Safety Systems - A Training Manual on Food Hygiene and the Hazard Analysis and Critical Control Point (HACCP) System. In.

FAO and WHO, 2012. Multicriteria-based ranking for risk management of foodborne parasites. In. Rome, Italy: FAO headquarters.

Feigenbaum A, Pinalli R, Giannetto M and Barlow S, in press. Reliability of the TTC approach: Learning from inclusion of pesticide active substances in the supporting database. Food and Chemical Toxicology.

Flari V and Chaudhry Q, 2012. Organisation of two horizon scanning international events on new sciences and technologies. In. York, UK: Fera.

Flari V, Chaudhry Q, Neslo R and Cooke R, 2011. Expert judgment based multi-criteria decision model to address uncertainties in risk assessment of nanotechnology-enabled food products. 13, 1813-1831.

Florig HK, Morgan MG, Morgan KM, Jenni KE, Fischhoff B, Fischbeck PS and DeKay ML, 2001. A Deliberative Method for Ranking Risks (I): Overview and Test Bed Development. Risk Analysis, 21, 913-913.

Food Safety Centre, 2010. Risk Ranger. In. Hobart, Australia.

Fosse J, Seegers H and Magras C, 2008a. Foodborne zoonoses due to meat: a quantitative approach for a comparative risk assessment applied to pig slaughtering in Europe. Veterinary Research, 39.

Fosse J, Seegers H and Magras C, 2008b. Prioritising the risk of foodborne zoonoses using a quantitative approach: application to foodborne bacterial hazards in pork and beef. Revue Scientifique Et Technique-Office International Des Epizooties, 27, 643-655.

French S, Maule J and Papamichail N, 2009. Decision Behaviour, Analysis and Support. Cambridge, UK: Cambridge University Press.

Frewer LJ, Kleter GA, Brennan M, Coles D, Fischer ARH, Houdebine LM, Mora C, Millar K and Salter B, in press. Genetically Modified Animals from Life-Science, Socio-Economic and Ethical Perspectives: Examining issues in an EU policy context. . New biotechnology.

Gadiel D, 2010. The economic cost of foodborne disease in New Zealand. . In (pp. 40 p.). Sydney, Australia: Applied Economics Pty Ltd.

Gamo M, Oka T and Nakanishi J, 2003. Ranking the risks of 12 major environmental pollutants that occur in Japan. Chemosphere, 53, 277-284.

Golan E, Buzby J, Crutchfield S, Frenzen PD, Kuchler F, Ralston K and Roberts T, . 2005. The value to consumers of reducing foodborne risks. In S. Hoffmann & M. R. Taylor (Eds.), Toward safer food. Perspectives on risk and priority setting. (pp. 129-158). Washington DC, USA: Resources for the future.

Gouin T, van Egmond R, Price OR and Hodges JEN, 2012. Prioritising chemicals used in personal care products in China for environmental risk assessment: Application of the RAIDAR model. Environmental Pollution, 165, 208-214.

Grace D, Monda J, Karanja N, Randolph TF and Kang'ethe EK, 2012. Participatory probabilistic assessment of the risk to human health associated with cryptosporidiosis from urban dairying in Dagoretti, Nairobi, Kenya. (Special Issue: Assessment and management of food-borne zoonoses in developing country cities.). Tropical animal health and production, 44, S33-S40.

Greim H and Reuter U, 2001. Classification of carcinogenic chemicals in the work area by the German MAK Commission: current examples for the new categories. Toxicology, 166, 11-23.

Guillier L, Danan C, Bergis H, Delignette-Muller ML, Granier S, Rudelle S, Beaufort A and Brisabois A, 2013. Use of quantitative microbial risk assessment when investigating foodborne illness outbreaks: The example of a monophasic Salmonella Typhimurium 4,5,12: I : - Outbreak implicating beef burgers. International Journal of Food Microbiology, 166, 471-478.

Haase A, Tentschert J and Luch A, 2012. Nanomaterials: a challenge for toxicological risk assessment? EXS, 101, 219-250.

Hammitt JK and Haninger K, 2007. Willingness to Pay for Food Safety: Sensitivity to Duration and Severity of Illness. American Journal of Agricultural Economics, 89, 1170-1175.

Haninger K and Hammitt JK, 2011. Diminishing willingness to pay per quality-adjusted life year: valuing acute foodborne illness. (Special Issue: Risk Regulation (Part 2): Risk Assessment and Economic Analysis). Risk Analysis, 31, 1363-1380.

Harrington JM, 1994. Research priorities in occupational medicine: A survey of United Kingdom medical opinion by the Delphi technique. Occupational and Environmental Medicine, 51, 289-294.

Havelaar AH, Galindo AV, Kurowicka D and Cooke RM, 2008. Attribution of foodborne pathogens using structured expert elicitation. Foodborne Pathogens & Disease, 5, 649-659.

Havelaar AH, van Rosse F, Bucura C, Toetenel MA, Haagsma JA, Kurowicka D, Heesterbeek JHAP, Speybroeck N, Langelaar MFM, van der Giessen JWB, Cooke RM and Braks MAH, 2010. Prioritizing emerging zoonoses in the Netherlands. PLoS ONE [Electronic Resource], 5, e13965.

Hoffmann S, 2010. Food safety policy and economics. A review of the literature. . In (pp. 39 p.). Washington DC, USA: Resources for the future.

Hoffmann S, Fischbeck P, Krupnick A and McWilliams M, 2008. Informing risk-mitigation priorities using uncertainty measures derived from heterogeneous expert panels: A demonstration using foodborne pathogens. Reliability Engineering and System Safety, 93, 687-698.

Hofstetter PH, J. K., 2002. Selecting human health metrics for environmental decision-support tools. Risk Analysis, 22, 965-983.

Hong T, Gurian PL, Huang Y and Haas CN, 2012. Prioritizing risks and uncertainties from intentional release of selected category A pathogens. PLoS ONE, 7.

Howard GA, M. F.; Teunis, P.; Mahmud, S. G.; Annette, D.; Deere, D., 2007. Disease burden estimation to support policy decision-making and research prioritization for arsenic mitigation. Journal of Water and Health, 5, 67-81.

Hristozov DR, Gottardo S, Critto A and Marcomini A, 2012. Risk assessment of engineered nanomaterials: a review of available data and approaches from a regulatory perspective. Nanotoxicology, 6, 880-898.

Humblet MF, Vandeputte S, Albert A, Gosset C, Kirschvink N, Haubruge E, Fecher-Bourgeois F, Pastoret PP and Saegerman C, 2012. Multidisciplinary and evidence-based method for prioritizing diseases of food-producing animals and zoonoses. Emerging Infectious Diseases, 18.

Juraske R, Antón A, Castells F and Huijbregts MAJ, 2007. PestScreen: A screening approach for scoring and ranking pesticides by their environmental and toxicological concern. Environment International, 33, 886-893.

Kemmeren JM, Mangen MJJ, van Duynhoven YTHP and Havelaar AH, 2006. Priority setting of foodborne pathogens: disease burden and costs of selected enteric pathogens. In Prioritering van voedsel overdraagbare pathogenen: ziektelast en ziektegebonden kosten: Rijksinstituut voor Volksgezondheid en Milieu RIVM.

Knight-Jones TJD, Mylrea GE and Kahn S, 2010. Animal production food safety: priority pathogens for standard setting by the World Organisation for Animal Health. Revue Scientifique Et Technique-Office International Des Epizooties, 29, 523-535.

Kokkinos P, Kozyra I, Lazic S, Bouwknegt M, Rutjes S, Willems K, Moloney R, de Roda Husman AM, Kaupke A, Legaki E, D'Agostino M, Cook N, Rzeżutka A, Petrovic T and Vantarakis A, 2012. Harmonised Investigation of the Occurrence of Human Enteric Viruses in the Leafy Green Vegetable Supply Chain in Three European Countries. 4, 179-191.

Krause G, 2008. How can infectious diseases be prioritized in public health? Embo Reports, 9, S22-S27.

Kruizinga AG, Briggs D, Crevel RWR, Knulst AC, Bosch LMCvd and Houben GF, 2008. Probabilistic risk assessment model for allergens in food: sensitivity analysis of the minimum eliciting dose and food consumption. Food and Chemical Toxicology, 46, 1437-1443.

Labite H and Cummins E, 2012. A Quantitative Approach for Ranking Human Health Risks from Pesticides in Irish Groundwater. Human and Ecological Risk Assessment, 18, 1156-1185.

Lachenmeier DW, Przybylski MC and Rehm J, 2012. Comparative risk assessment of carcinogens in alcoholic beverages using the margin of exposure approach. International Journal of Cancer, 131, E995-E1003.

Lake RJ, Cressey PJ, Campbell DM and Oakley E, 2010. Risk Ranking for Foodborne Microbial Hazards in New Zealand: Burden of Disease Estimates. Risk Analysis, 30, 743-752.

Lenartowicz P and Michie N, 2002a. Risk-based sampling of food. Vol 1. Risk assessment for sampling. In Public Analyst Service Ltd (Ed.), (pp. 16 p.): Food Standards Agency.

Lenartowicz P and Michie N, 2002b. Risk-based sampling of food. Vol 2. Background and support. In Public Analyst Service Ltd (Ed.), (pp. 32 p.): Food Standards Agency.

Lichtenstein S and Slovic Pe, 2006. The construction of preference Cambridge, UK: Cambridge University Press.

Lim SSV T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Amann M, Anderson HR, Andrews KG, Aryee M, Atkinson C, Bacchus LJ, Bahalim AN, Balakrishnan K, Balmes J, Barker-Collo S, Baxter A, Bell ML, Blore JD, Blyth F, Bonner C, Borges G, Bourne R, Boussinesq M, Brauer M, Brooks P, Bruce NG, Brunekreef B, Bryan-Hancock C, Bucello C, Buchbinder R, Bull F, Burnett RT, Byers TE, Calabria B, Carapetis J, Carnahan E, Chafe Z, Charlson F, Chen H, Chen JS, Cheng ATA, Child JC, Cohen A, Colson KE, Cowie BC, Darby S, Darling S, Davis A, Degenhardt L, Dentener F, Des Jarlais DC, Devries K, Dherani M, Ding EL, Dorsey ER, Driscoll T, Edmond K, Ali SE, Engell RE, Erwin PJ, Fahimi S, Falder G, Farzadfar F, Ferrari A, Finucane MM, Flaxman S, Fowkes FGR, Freedman G, Freeman MK, Gakidou E, Ghosh S, Giovannucci E, Gmel G, Graham K, Grainger R, Grant B, Gunnell D, Gutierrez HR, Hall W, Hoek HW, Hogan A, Hosgood Iii HD, Hoy D, Hu H, Hubbell BJ, Hutchings SJ, Ibeanusi SE, Jacklyn GL, Jasrasaria R, Jonas JB, Kan H, Kanis JA, Kassebaum N, Kawakami N, Khang YH, Khatibzadeh S, Khoo JP, Kok C, Laden F, Lalloo R, Lan Q, Lathlean T, Leasher JL, Leigh J, Li Y, Lin JK, Lipshultz SE, London S, Lozano R, Lu Y, Mak J, Malekzadeh R, Mallinger L, Marcenes W, March L, Marks R, Martin R, McGale P, McGrath J, Mehta S, Mensah GA, Merriman TR, Micha R, Michaud C, Mishra V, Hanafiah KM, Mokdad AA, Morawska L, Mozaffarian D, Murphy T, Naghavi M, Neal B, Nelson PK, Nolla JM, Norman R, Olives C, Omer SB, Orchard J, Osborne R, Ostro B, Page A, Pandey KD, Parry CDH, Passmore E, Patra J, Pearce N, Pelizzari PM, Petzold M, Phillips MR, Pope D, Pope Iii CA, Powles J, Rao M, Razavi H, Rehfuess EA, Rehm JT, Ritz B, Rivara FP, Roberts T, Robinson C, Rodriguez-Portales JA, Romieu I, Room R, Rosenfeld LC, Roy A, Rushton L,

Salomon JA, Sampson U, Sanchez-Riera L, Sanman E, Sapkota A, Seedat S, Shi P, Shield K, Shivakoti R, Singh GM, Sleet DA, Smith E, Smith KR, Stapelberg NJC, Steenland K, Stöckl H, Stovner LJ, Straif K, Straney L, Thurston GD, Tran JH, Van Dingenen R, Van Donkelaar A, Veerman JL, Vijayakumar L, Weintraub R, Weissman MM, White RA, Whiteford H, Wiersma ST, Wilkinson JD, Williams HC, Williams W, Wilson N, Woolf AD, Yip P, Zielinski JM, Lopez AD, Murray CJL and Ezzati M. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. The Lancet, 380, 2224-2260.

Linkov I, Bates ME, Trump BD, Seager TP, Chappell MA and Keisler JM, 2013. For nanotechnology decisions, use decision analysis. Nano Today, 8, 5-10.

Madsen CB, Hattersley S, Buck J, Gendel SM, Houben GF, Hourihane JO, Mackie A, Mills ENC, Nørhede P, Taylor SL and Crevel RWR, 2009. Approaches to risk assessment in food allergy: Report from a workshop "developing a framework for assessing the risk from allergenic foods". Food and Chemical Toxicology, 47, 480-489.

Malorny B, Löfström C, Wagner M, Krämer N and Hoorfar J, 2008. Enumeration of Salmonella bacteria in food and feed samples by real-time PCR for quantitative microbial risk assessment. Applied and Environmental Microbiology, 74, 1299-1304.

Mangen MJ, Bouwknegt M, Friesema IHM, Kortbeek LM, van Pelt W and Havelaar AH, 2013. Disease burden and cost-of-illness of food-related pathogens in the Netherlands. In (pp. 95 p.). Bilthoven, the Netherlands: National Institute of Public health and the Environment (RIVM).

Mangen MJJ, Batz MB, Käsbohrer A, Hald T, Morris Jr JG, Taylor M and Havelaar AH, 2010. Integrated approaches for the public health prioritization of foodborne and zoonotic pathogens. Risk Analysis, 30, 782-797.

Mangen MJJ, De Wit GA and Havelaar AH, 2007. Economic analysis of Campylobacter control in the Dutch broiler meat chain. (Special Issue: Economic measures of food safety interventions.). Agribusiness, 23, 173-192.

Mangen MJJ, Plass D and Kretzschmar MEE, 2014. Estimating the current and future burden of communicable diseases in the European Union and EEA/EFTA. In RIVM Report 210474001/2014.

Mara DD, Sleigh PA, Blumenthal UJ and Carr RM, 2007. Health risks in wastewater irrigation: Comparing estimates from quantitative microbial risk analyses and epidemiological studies. Journal of Water and Health, 5, 39-50.

Maunula L, Kaupke A, Vasickova P, Söderberg K, Kozyra I, Lazic S, van der Poel WHM, Bouwknegt M, Rutjes S, Willems KA, Moloney R, D'Agostino M, de Roda Husman AM, von Bonsdorff C-H, Rzeżutka A, Pavlik I, Petrovic T and Cook N, 2013. Tracing enteric viruses in the European berry fruit supply chain. International Journal of Food Microbiology, 167, 177-185.

McSpirit S and Reid C, 2010. Residents' Perceptions of Tap Water and Decisions to Purchase Bottled Water: A Survey Analysis from the Appalachian, Big Sandy Coal Mining Region of West Virginia. Society & Natural Resources, 24, 511-520.

Micha R, Kalantarian S, Wirojratana P, Byers T, Danaei G, Elmadfa I, Ding E, Giovannucci E, Powles J, Smith-Warner S, Ezzati M and Mozaffarian D, 2012. Estimating the global and regional burden of suboptimal nutrition on chronic disease: Methods and inputs to the analysis. European Journal of Clinical Nutrition, 66, 119-129.

Miller GY, Liu X, McNamara PE and Barber DA, 2005. Influence of *Salmonella* in Pigs Preharvest and during Pork Processing on Human Health Costs and Risks from Pork. Journal of Food Protection, 68, 1788-1798.

Mitchell J, Arnot JA, Jolliet O, Georgopoulos PG, Isukapalli S, Dasgupta S, Pandian M, Wambaugh J, Egeghy P, Cohen Hubal EA and Vallero DA, 2013. Comparison of modeling approaches to prioritize chemicals based on estimates of exposure and exposure potential. Science of the Total Environment, 458-460, 555-567.

Mitchell J, Pabon N, Collier ZA, Egeghy PP, Cohen-Hubal E, Linkov I and Vallero DA, 2013. A Decision Analytic Approach to Exposure-Based Chemical Prioritization. PLoS ONE, 8.

Moffet J, 1996. Environmental priority setting based on comparative risk and public input. Canadian Public Administration, 39, 362-385.

Morgan K, 2005. Development of a preliminary framework for informing the risk analysis and risk management of nanoparticles. Risk Analysis, 25, 1621-1635.

Mørkbak M and Nordström J, 2009. The Impact of Information on Consumer Preferences for Different Animal Food Production Methods. Journal of Consumer Policy, 32, 313-331.

Mørkbak MR, Christensen T and Gyrd-Hansen D, 2011. Consumers' willingness to pay for safer meat depends on the risk reduction methods - A Danish case study on Salmonella risk in minced pork. Food Control, 22, 445-451.

Neslo R, 2011. Discrete decisions with model validation using probabilistic inversion. University of Delft, Delft, The Netherlands.

Neslo RE and Cooke RM, 2011. Modeling and validating stakeholder preferences with probabilistic inversion. Applied Stochastic Models in Business and Industry, 27, 115-130.

Newsome RT, N.; Paoli, G. M.; Jaykus, L. A.; Tompkin, B.; Miliotis, M.; Ruthman, T.; Hartnett, E.; Busta, F. F.; Petersen, B.; Shank, F.; McEntire, J.; Hotchkiss, J.; Wagner, M.; Schaffner, D. W., 2009. Development of a risk-ranking framework to evaluate potential high-threat microorganisms, toxins, and chemicals in food. Journal of Food Science, 74, R39-R45.

O'Brien NJ and Cummins EJ, 2011. A Risk Assessment Framework for Assessing Metallic Nanomaterials of Environmental Concern: Aquatic Exposure and Behavior. Risk Analysis, 31, 706-726.

Oldenkamp R, Huijbregts MAJ, Hollander A, Versporten A, Goossens H and Ragas AMJ, 2013. Spatially explicit prioritization of human antibiotics and antineoplastics in Europe. Environment International, 51, 13-26.

Pennington D, Crettaz P, Tauxe A, Rhomberg L, Brand K and Jolliet O, 2002. Assessing human health response in life cycle assessment using ED10s and DALYs: Part 2 - Noncancer effects. Risk Analysis, 22, 947-963.

Penrose LJ, Thwaite WG and Bower CC, 1994. Rating index as a basis for decision making on pesticide use reduction and for accreditation of fruit produced under integrated pest management. Crop Protection, 13, 146-152.

Polinder S, Haagsma J, Stein C and Havelaar A, 2012. Systematic review of general burden of disease studies using disability-adjusted life years. 10, 1-15.

Pouillot R and Lubran MB, 2011. Predictive microbiology models vs. modeling microbial growth within Listeria monocytogenes risk assessment: What parameters matter and why. Food Microbiology, 28, 720-726.

Presi P, Stärk K, Knopf L, Breidenbach E, Sanaa M, Frey J and Regula G, 2008. Efficiency of risk-based vs. random sampling for the monitoring of tetracycline residues in slaughtered calves in Switzerland. Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 25, 566-573.

Ranke J and Jastorff B, 2000. Multidimensional risk analysis of antifouling biocides. Environmental Science and Pollution Research, 7, 105-114.

Reist M, Jemmi T and Stärk KDC, 2012. Policy-driven development of cost-effective, risk-based surveillance strategies. Preventive Veterinary Medicine, 105, 176-184.

Rietjens IMCM, Slob W, Galli C and Silano V, 2008. Risk assessment of botanicals and botanical preparations intended for use in food and food supplements: Emerging issues. Toxicology Letters, 180, 131-136.

Robichaud CO, Tanzie D, Weilenmann U and Wiesner MR, 2005. Relative risk analysis of several manufactured nanomaterials: An insurance industry context. Environmental Science and Technology, 39, 8985-8994.

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

Rowe G and Frewer LJ, 2005. A Typology of Public Engagement Mechanisms. Science, Technology & Human Values, 30, 251-290.

Ruzante JM, Davidson VJ, Caswell J, Fazil A, Cranfield JAL, Henson SJ, Anders SM, Schmidt C and Farber JM, 2010. A multifactorial risk prioritization framework for foodborne pathogens. Risk Analysis, 30, 724-742.

Safe SH, 1998. Hazard and risk assessment of chemical mixtures using the toxic equivalency factor approach. Environmental Health Perspectives, 106, 1051-1058.

Sailaukhanuly Y, Zhakupbekova A, Amutova F and Carlsen L, 2013. On the ranking of chemicals based on their PBT characteristics: Comparison of different ranking methodologies using selected POPs as an illustrative example. Chemosphere, 90, 112-117.

Schmidt K, Höflich C, Bruch M, Entzian K, Horn P, Kacholdt A, Kragl U, Leinweber P, Mikschofsky H, Mönkemeyer W, Mohr E, Neubauer K, Schlichting A, Schmidtke J, Steinmann A, Struzyna-Schulze C, Wilhelm R, Zeyner A, Ziegler A and Broer I, 2011. BioOK - A comprehensive system for analysis and risk assessment of genetically modified plants. Journal fur Kulturpflanzen, 63, 232-248.

Schwarzinger M, Mohamed MK, Gad RR, Dewedar S, Fontanet A, Carrat F and Luchini S, 2010. Risk perception and priority setting for intervention among hepatitis C virus and environmental risks: A cross-sectional survey in the Cairo community. BMC Public Health, 10.

Sinclair CJ, Boxall ABA, Parsons SA and Thomas MR, 2006. Prioritization of pesticide environmental transformation products in drinking water supplies. (Special issue: Emerging contaminants.). Environmental Science & Technology, 40, 7283-7289.

Son YK, Yoon CG, Rhee HP and Lee SJ, 2013. A review on microbial and toxic risk analysis procedure for reclaimed wastewater irrigation on paddy rice field proposed for South Korea. 11, 543-550.

Sorensen PB, Thomsen M, Assmuth T, Grieger KB and Baun A, 2010. Conscious worst case definition for risk assessment, part I A knowledge mapping approach for defining most critical risk factors in integrative risk management of chemicals and nanomaterials. Science of the Total Environment, 408, 3852-3859.

Spanjersberg MQI, Kruizinga AG, Rennen MAJ and Houben GF, 2007. Risk assessment and food allergy: the probabilistic model applied to allergens. Food and Chemical Toxicology, 45, 49-54.

Stella P, Cerf O, Hugas M, Koutsoumanis KP, Nguyen-The C, Sofos JN, Valero A and Zwietering MH, 2013. Ranking the microbiological safety of foods: a new tool and its application to composite products. Trends in Food Science & Technology, 33, 124-138.

Taxell P, Engström K, Tuovila J, Söderström M, Kiljunen H, Vanninen P and Santonen T, 2013. Methodology for national risk analysis and prioritization of toxic industrial chemicals. Journal of Toxicology and Environmental Health - Part A: Current Issues, 76, 690-700.

Teck SJ, Halpern BS, Kappel CV, Micheli F, Selkoe KA, Crain CM, Martone R, Shearer C, Arvai J, Fischhoff B, Murray G, Neslo R and Cooke R, 2010. Using expert judgment to estimate marine ecosystem vulnerability in the California Current. Ecological Applications, 20, 1402-1416.

Tran A, Ippoliti C, Balenghien T, Conte A, Gely M, Calistri P, Goffredo M, Baldet T and Chevalier V, 2013. A Geographical Information System-Based Multicriteria Evaluation to Map Areas at Risk for Rift Valley Fever Vector-Borne Transmission in Italy. Transboundary and Emerging Diseases, 60, 14-23.

Travisi CM, Nijkamp P, Vighi M and Giacomelli P, 2006. Managing pesticide risks for non-target ecosystems with pesticide risk indicators: A multi-criteria approach. International Journal of Environmental Technology and Management, 6, 141-162.

Umasunthar T, Leonardi-Bee J, Hodes M, Turner PJ, Gore C, Habibi P, Warner JO and Boyle RJ, 2013. Incidence of fatal food anaphylaxis in people with food allergy: a systematic review and meta-analysis. Clinical & Experimental Allergy, 43, 1333-1341.

Valcke M, Chaverri F, Monge P, Bravo V, Mergler D, Partanen T and Wesseling C, 2005. Pesticide prioritization for a case-control study on childhood leukemia in Costa Rica: A simple stepwise approach. Environmental Research, 97, 335-347.

Van Asselt ED, Sterrenburg P, Noordam MY and Van der Fels-Klerx HJ, 2012. Overview of available methods for Risk Based Control within the European Union. Trends in Food Science & Technology, 23, 51-58.

van Asselt ED, van der Spiegel M, Noordam MY, Pikkemaat MG and van der Fels-Klerx HJ, 2013. Risk ranking of chemical hazards in food-A case study on antibiotics in the Netherlands. Food Research International, 54, 1636-1642.

Van der Fels-Klerx HJ, Goossens LHJ, Saatkamp HW and Horst SHS, 2002. Elicitation of Quantitative Data from a Heterogeneous Expert Panel: Formal Process and Application in Animal Health. Risk Analysis, 22, 67-81.

van Dijk H, Fischer AR and Frewer LJ, 2011. Consumer Responses to Integrated Risk-Benefit Information Associated with the Consumption of Food. Risk Analysis, 31, 429-439.

van Kreijl CF, Knaap AGAC, Busch MCM, Havelaar AH, Kramers PGN, Kromhout D, van Leeuwen FXR, van Leent-Loenen HMJA, Ocké MC and Verkleij H, 2004. Ons eten gemeten - Gezonde voeding en veilig voedsel in Nederland. In (pp. 364 p.). Bilthoven, the Netherlands: National Institute of Public Health and the Environment (RIVM).

Van Lier EA, Havelaar AH and Nanda A, 2007. The burden of infectious diseases in Europe: a pilot study. Euro Surveill, 12, E3-4.

VRC, 2010. Annual Report on Surveillance for Veterinary Residues in Food in the UK 2010. In (pp. 51 p.). Surrey, UK: Veterinary Residues Committee.

Wambaugh JF, Setzer RW, Reif DM, Gangwal S, Mitchell-Blackwood J, Arnot JA, Joliet O, Frame A, Rabinowitz J, Knudsen TB, Judson RS, Egeghy P, Vallero D and Cohen Hubal EA, 2013. High-throughput models for exposure-based chemical prioritization in the ExpoCast project. Environmental Science and Technology, 47, 8479-8488.

Whiteside M, Mineau P, Morrison C and Knopper LD, 2008. Comparison of a score-based approach with risk-based ranking of in-use agricultural pesticides in Canada to aquatic receptors. Integrated Environmental Assessment and Management, 4, 215-236.

WHO, 2009. Principles and methods for the risk assessment of chemicals in food. In Environmental Health Criteria; 2009. (240):lxix + 685 pp. many ref. Geneva: World Health Organization.

Williams MS and Ebel ED, 2012. Methods for fitting the Poisson-lognormal distribution to microbial testing data. Food Control, 27, 73-80.

Xiao S, An W, Chen Z, Zhang D, Yu J and Yang M, 2012. The burden of drinking water-associated cryptosporidiosis in China: The large contribution of the immunodeficient population identified by quantitative microbial risk assessment. Water Research, 46, 4272-4280.

Zalk DM, Paik SY and Swuste P, 2009. Evaluating the Control Banding Nanotool: A qualitative risk assessment method for controlling nanoparticle exposures. Journal of Nanoparticle Research, 11, 1685-1704.

Zuin S, Micheletti C, Critto A, Pojana G, Johnston H, Stone V, Tran L and Marcomini A, 2011. Weight of Evidence approach for the relative hazard ranking of nanomaterials. Nanotoxicology, 5, 445-458.

APPENDIX/APPENDICES

Annex I. Approaches have been already employed by EFSA to rank foodborne health risks due to contamination by biological hazards.

Table A1. Ranking approaches employed by the BIOHAZ EFSA scientific panel to evaluate either attribution of hazards to foods and/or foodborne health risks due to contamination by a biological hazard are presented in terms of main data inputs, integration and analysis of data, and outputs of the approach.

Source	Main characteristics of the approach			
	Source of main data inputs Top-down or bottom- up	Description of approach	Quantitative / Qualitative	Integration and analysis of data; outputs
EFSA 2012: Scientific Opinion on public health risks represented by certain composite products containing food of animal origin. EFSA Journal, 10(5): 2662, 132 pp.[4]	Top-down	2D matrix	Qualitative	Scoring *via* expert judgment – categories of food/hazards combinations
	Bottom-up	Decision tree	Qualitative	Flow of decision tree *via* expert judgment – risk categories
EFSA 2012: Scientific Opinion on the public health hazards to be covered by inspection of meat from poultry. EFSA Journal, 10(6): 2741[5]	Bottom-up	Decision tree	Qualitative	Scoring *via* expert judgment – risk categories

[4] http://www.efsa.europa.eu/en/efsajournal/doc/2662.pdf
[5] http://www.efsa.europa.eu/en/efsajournal/doc/2741.pdf

88

EFSA 2011: Scientific Opinion on the public health hazards to be covered by inspection of meat (swine). EFSA Journal, 9(10):2351, 198pp.	Top-down	Classification[6] tree	Quantitative	Aggregation of diverse data inputs *via* expert judgment – risk classes representing probability ranges
EFSA 2011: Scientific Opinion on a quantitative estimation of the public health impact of setting a new target for the reduction of Salmonella in broilers. EFSA Journal, 9(7):2106, 94 pp.[7]	Top-down	Deterministic source attribution model[8]	Quantitative	Proportions of contribution of food sources to human campylobacteriosis
EFSA 2012: Scientific Opinion on the public health hazards to be covered by inspection of meat from poultry. EFSA Journal 2012, 10(6): 2741[9]	Top-down	Deterministic source attribution model[10]	Quantitative	Proportions of contribution of food sources to human salmonellosis
EFSA 2010: Scientific Opinion on risk assessment of parasites in fishery products. EFSA Journal, 8(4): 1543, 91 pp.[11]	Bottom-up	Multi criteria approach	Qualitative	It is assumed that aggregation was *via* expert judgment – risk categories (incl. not known)
EFSA 2008: Scientific Opinion of the Panel on Biological Hazards on foodborne	Bottom-up	Risk pathway exposure model –	Quantitative	Aggregation *via* expert judgment, *i.e.* interpretation of data or subjective probabilities – risk classes

---

[6] Therefore probability ranges were used as classes of risk

[7] http://www.efsa.europa.eu/en/efsajournal/doc/2106.pdf

[8] BT-SAM model: XXX; TT-SAM model in Hald et al., 2012: EFSA Supporting publication 2012: EN-259 (No 20 in our excel spreadsheet).

[9] http://www.efsa.europa.eu/en/efsajournal/doc/2741.pdf

[10] Employing microbial subtyping (MLST).

[11] http://www.efsa.europa.eu/en/search/doc/1543.pdf

antimicrobial resistance as a biological hazard. EFSA Journal. 765, 1-87.[12]		probability entries at high level		representing probability ranges
EFSA 2007: Scientific Opinion of the Panel on Biological Hazards on the revision of the Geographical BSE risk assessment (GBR) methodology. EFSA Journal. 463, 1-35.[13]	Bottom-up	Stability model of BSE-agent	Semi-quantitative	Aggregation *via* modelling – leading to a categorisation representing ranges of reproduction ratio
SCVMPH 2003: Opinion of the Scientific Committee on Veterinary Measures relating to Public Health on Salmonella in foodstuffs. 14 April 2003, 1-65.[14]	Top-down and Bottom-up	Overall assessment of available information	Qualitative	Aggregation *via* expert judgment – leading to risk categories
SCVMPH 2003: Opinion of the Scientific Committee on Veterinary Measures relating to Public Health on verotoxigenic E. coli (VTEC) in foodstuffs. 21-22 January 2003, 1-65.[15]	Top-down and Bottom-up	Overall assessment of available information	Qualitative	Aggregation *via* expert judgment – leading to risk categories

---

[12] http://www.efsa.europa.eu/en/efsajournal/doc/765.pdf
[13] http://www.efsa.europa.eu/en/efsajournal/doc/463.pdf
[14] http://ec.europa.eu/food/fs/sc/scv/out66_en.pdf
[15] http://ec.europa.eu/food/fs/sc/scv/out58_en.pdf

Annex II. Outcome of the decision tool for the selected case studies

Table A2. Outcome of the decision tool for the case study on mycotoxins

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/ Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled
METHOD COMMENTS			Possibly appropiate method	Possibly appropiate method				Possibly appropiate method			
			Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation	Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation				Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation User-friendly tools may be availabe			

Table A3. Outcome of the decision tool for the case study on PA, part I

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/ Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled
	Data not available										
METHOD COMMENTS											

Table A4. Outcome of the decision tool for the case study on PA, part II

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/ Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available
METHOD COMMENTS			**Possibly appropiate method**	**Possibly appropiate method**				**Possibly appropiate method**	**Possibly appropiate method**	**Possibly appropiate method**	**Possibly appropiate method**
			Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation	Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation				Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation User-friendly tools may be availabe	Resources: High Moderate Low Accuracy: Qualitative Outputs: Graph	Resources: High Moderate Low Accuracy: Qualitative Outputs: Decision Tree	Resources: High Moderate Low Accuracy: Qualitative Uncertainty included Outputs: Tabulation

Table A5. Outcome of the decision tool for the case study on botanicals

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/ Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites fulfilled Data not available	Prerequsites not fulfilled	Prerequsites fulfilled Data not available	Prerequsites fulfilled Data not available	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available
METHOD COMMENTS								**Possibly appropiate method**	**Possibly appropiate method**	**Possibly appropiate method**	**Possibly appropiate method**
								Resources: High Moderate	Resources: High Moderate Low	Resources: High Moderate Low	Resources: High Moderate Low
								Accuracy: Semi-quantitative	Accuracy: Qualitative	Accuracy: Qualitative	Accuracy: Qualitative
								Uncertainty included			Uncertainty included
								Outputs: Tabulation	Outputs: Graph	Outputs: Decision Tree	Outputs: Tabulation
								User-friendly tools may be availabe			

Table A6. Outcome of the decision tool for the case study on zoonoses, part I

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled
METHOD COMMENTS					**Possibly appropiate method**	**Possibly appropiate method**		**Possibly appropiate method**			
					Resources: Moderate	Resources: High Moderate		Resources: High Moderate			
					Accuracy: Quantitative Semi-quantitative Uncertainty included	Accuracy: Quantitative Semi-quantitative Uncertainty included		Accuracy: Semi-quantitative Uncertainty included			
					Outputs: Graph Tabulation	Outputs: Graph Tabulation		Outputs: Tabulation			
					User-friendly tools may be availabe	User-friendly tools may be availabe		User-friendly tools may be availabe			

Table A7. Outcome of the decision tool for the case study on zoonoses, part II

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available
METHOD COMMENTS			Possibly appropiate method	Possibly appropiate method		Possibly appropiate method		Possibly appropiate method			Possibly appropiate method
			Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation	Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation		Resources: High Moderate Accuracy: Quantitative Semi-quantitative Uncertainty included Outputs: Graph Tabulation User-friendly tools may be availabe		Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation User-friendly tools may be availabe			Resources: High Moderate Low Accuracy: Qualitative Uncertainty included Outputs: Tabulation

Table A8. Outcome of the decision tool for the case study on viruses

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/ Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled
METHOD COMMENTS								**Possibly appropiate method**			
								Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation User-friendly tools may be availabe			

Table A9. Outcome of the decision tool for the case study on nutritional hazards

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/ Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites fulfilled	Prerequsites fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled
	Data available	Data available				Data available					
METHOD COMMENTS	**Possibly appropiate method**	**Possibly appropiate method**				**Possibly appropiate method**					
	Resources: High	Resources: High				Resources: High Moderate					
	Accuracy: Quantitative	Accuracy: Quantitative				Accuracy: Quantitative Semi-quantitative Uncertainty included					
	Uncertainty included	Uncertainty included									
	Outputs: Graph Tabulation	Outputs: Graph Tabulation				Outputs: Graph Tabulation					
						User-friendly tools may be availabe					

Table A10. Outcome of the decision tool for the case study on microbiological, chemical and nutritional hazards

EVALUATION	Risk Assessment'	Comparative Risk Assessment'	Ratio (Exposure/ Effect)'	Scoring method	Cost of Illness'	DALY/QALY'	WTP'	MCDA'	Risk Matrix'	Flow charts /Decision trees'	Expert Synthesis'
	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites fulfilled Data available	Prerequsites not fulfilled	Prerequsites not fulfilled	Prerequsites not fulfilled
METHOD COMMENTS						**Possibly appropiate method**		**Possibly appropiate method**			
						Resources: High Moderate Accuracy: Quantitative Semi-quantitative Uncertainty included Outputs: Graph Tabulation User-friendly tools may be availabe		Resources: High Moderate Accuracy: Semi-quantitative Uncertainty included Outputs: Tabulation User-friendly tools may be availabe			

Annex III. Proposed design of an expert workshop to ranking botanicals

The proposed method for designing an expert workshop for ranking the risks related to botanicals broadly follows the Carnegie Mellon approach for risk ranking (Florig et al., 2001) though differs in that deliberation and ranking are performed by experts rather than lay people, due to the need to consider complex and incomplete scientific information. It loosely follows the procedure mentioned in Florig et al. (2001):

- Step A: Define the risks to be ranked,

- Step B: Identify the risk attributes that should be considered,

- Step C: Describe the risks in terms of the attributes in risk summary sheets,

- Step D: Select participants. Conduct a workshop at which risk ranking is performed,

- Step E: Describe the issues identified and the resulting rankings.

Step A: Preliminary workshop to select botanicals for consideration

The EFSA Compendium identifies around 900 plant species with potentially adverse impacts on human safety, probably ranging from marginal to more severe impacts. Available evidence would probably allow a presumption of safety to be made for a proportion of these which therefore would not require further consideration. The number which could be deliberated upon within reasonable time constraints without inducing respondent overload is much smaller. For example the Unfinished Business project (US EPA, 1987) ranked 31 environmental hazards. Application of the Carnegie-Mellon method to hypothetical risks at a school addressed 22 hazards (Florig et al., 2001), and in a multi criteria based ranking of foodborne parasites, 24 parasites out of a possible total of 95 were selected for consideration over 4 days (FAO/WHO, 2012). Therefore, the aim of Stage 1 is to identify the 10 to 20 botanicals with the potential to cause the highest negative public health impacts (or smallest net benefit). To ensure a consistent approach, a series of criteria should be identified by which botanicals will be selected. The criteria proposed are:

- Evidence that the substance is potentially harmful (e.g., contains an ingredient which is toxic or carcinogenic), or it is banned in one or more countries, or there is anecdotal or other evidence of harmful effects in humans, or there is similarity to another substance about which there are human health concerns,

- Potential use is at higher dosage than occurs from historical use,

- It has the potential for much wider future use in the population (e.g. if there is commercial interest in producing it).

These criteria are derived from those identified by EFSA (2009) for use in priority setting in safety assessment for botanicals. To reduce the identified botanicals to a manageable number it will probably be necessary to introduce additional criteria, such as severity and likelihood of negative impacts. By screening within this framework, substances which are not harmful at normal use, with low toxic potential, and which are consumed at low doses will be identified and not be considered further.

Expert judgement is required for such a process, and a workshop would be the preferred means of obtaining it. The large scale of this operation implies a need for separate teams which each examine a proportion of the total substances (with some overlap for validation purposes), and a scoring system or series of defined judgements. This could follow WCRF/AICR criteria for grading evidence: convincing, probable, limited-suggestive, limited-no conclusion or substantial effect on risk unlikely.

Steps B and C: Identify the risk attributes that should be considered and summarise the available information.

Here, it is proposed that there is some deviation from the Carnegie-Mellon approach and the evidence presented to experts at the deliberative workshop is in written form and follows a similar format to that used in the ESCO case studies, namely short literature reviews which summarise what is known about the action and safety of the substance. (This is much more detailed than the information typically presented in workshops for lay people.)

Steps D and E: select experts and conduct workshop at which experts perform a risk ranking of selected botanicals. Documentation of results.

Experts are required as the tasks involved are understanding, interpreting and evaluating the available information and making comparisons between substances on the basis of incomplete information. To deal with the multi-faceted nature of the problem, experts should be drawn from a variety of disciplinary backgrounds.

The ranking produced by each expert, or collectively, should be regarded as uncertain but nevertheless the best as can be managed for the current state of knowledge. Uncertainty is reflected by the degree of consensus or disagreement. Of equal, if not greater importance is the content of the deliberations which can reveal, for example, why a ranking was given; which factors are critical; how the problem of missing information is dealt with; the degree of confidence in the outcome, and reasons for non-consensus. Benefits and risks also need to be communicated separately as they may accrue to different constituencies (Tijhuis et al., 2012). This approach does not produce a straightforward ranking that is easy to operationalise. Rather it produces detailed and complex qualitative information ('thick description') which provides deeper insights.

Annex IV. Two possible MCDA approaches for risk prioritization of foodborne viruses in berry fruits and leafy green vegetables

Alternative 1: Risk ranking using the MCDA approach developed by EFSA BIOHAZ panel

EFSA BIOHAZ panel opinion (2013) suggests a MCDA to address risk ranking of food/pathogen combinations of food of non-animal origin. The model developed by the EFSA BIOHAZ panel was principally based on the FSA Risk Ranking tool (P^3-ARRT) developed by Anderson et al. (2011). The criteria employed by the EFSA BIOHAZ panel for the 2013 opinion were:

- Strength of associations between food and pathogen based on the foodborne outbreak from EU Zoonoses Monitoring Data

- Incidence of illness

- Burden of disease

- Dose response relationship

- Consumption

- Prevalence of contamination

- Pathogen growth potential during shelf life

The EFSA BIOHAZ panel grouped available data for each criterion into scoring categories, which were defined and assigned a numerical, ordinal score. The criteria, the definition of the scoring categories and the data applied are described in detail in the opinion (EFSA, 2013). Finally, for each food/pathogen combination, the scores of the seven criteria were summed to give a total final risk score, thereby providing a ranking of all combinations. Based on these criteria, a customised list of criteria to address the particular risk ranking exercise on foodborne viruses in berry fruits and leafy green vegetables was devised:

- Strength of associations between food commodity and virus based on outbreak data

- Strength of associations between food commodity and virus based on prevalence data[16]

- Incidence of illness[17]

- Burden of disease

- Consumption of the particular food in the food/virus combination

- Secondary spread of virus

The dose response relationship criterion suggested by EFSA (2013) was not included in the suggested list of criteria for foodborne viruses in berry fruits and leafy green vegetables, because there is no information on dose response relationships for either HAV or HEV and it is highly unlikely that this information would be acquired in the future. Moreover, the infectious dose of each virus, concerned in this case study, was expected to be very low, for example, within the range of 10-100 virus particles[18].

Once new criteria are suggested for the a risk ranking of foodborne viruses in berry fruits and leafy green vegetables combinations, one would need to devise scoring categories for the data for each one, most probably via a remote questionnaire to experts and/or via structured expert workshops. Alternatively, one could apply the MCDA approach shown in the following section, where each discrete criterion is described via a measurable unit.

Alternative 2: Risk ranking using the MCDA approach based on inferred weights of criteria

A much more recent approach involves the application of probabilistic inversion in MCDA modelling which allows for inferring the weights of each criterion by each expert (instead of eliciting directly the weights from the experts). The method has been applied in food safety related problems (Flari et al., 2011; Havelaar et al., 2010). Havelaar et al. (2010) compared the weights derived by probabilistic inversion with the simple ranking method (as proposed by (Krause, 2008), and found not significant correlation between the two methods. One would need to follow a number of steps, illustrated graphically in Figure A1 below, in order to implement it.

---

[16] Prevalence data for foodborne viruses is hindered by lack of data (EFSA 2013; EFSA 2014a). It is possible that low level contamination may result in numerous sporadic cases which could go unnoticed within the community, and thus not reported further.

[17] This criterion would be applicable for all viruses considered in this case study, *i.e.* Norovirus, HAV, HEV, but for the moment incidence of illness data exist only for HEV.

[18] Detailed data on Infectious Dose 50 (ID50) and probability of becoming infected are available only for Norovirus (in Teunis *et al.*, 2008).
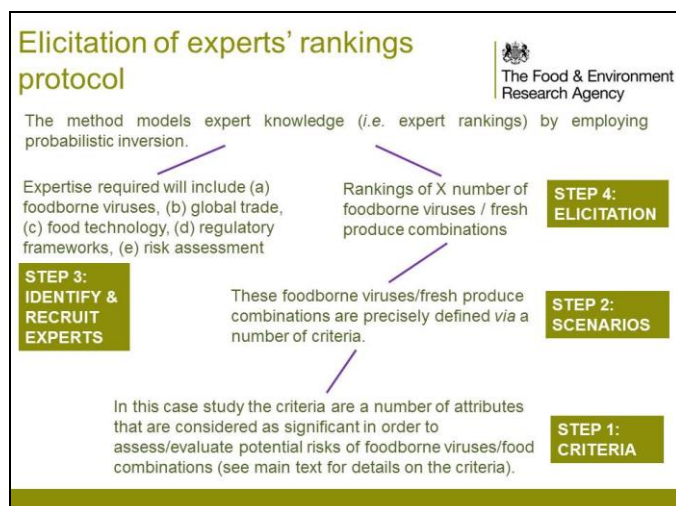
Figure A1: Illustration of steps required to implement the MCDA with probabilistic inversion approach. A risk analyst with expertise in this particular approach is essential for the customised development of the approach and for building the necessary mathematical models. In particular:

-   Identification of discrete, measurable criteria and measurement units for each criterion: a list of potentially suitable criteria for ranking foodborne viruses/berry fruits and leafy green vegetables combination is cited above, and each criterion was characterised by a measurable unit as shown in Table A11.

Table A11. List of potentially suitable criteria (and their units) for ranking foodborne viruses/berry fruits and leafy green vegetables.

Criterion		Measurable unit
C1	Strength of associations between food and virus based on outbreak data	No of outbreaks in total [19] per virus per food commodity[20]
C2	Strength of associations between food and virus based on prevalence data	Percentage of virus positive samples per food commodity
C3	Incidence of illness	Annual incidences in the community (e.g. EU)
C4	Burden of disease	DALYs per food commodity[21]
C5	Consumption of the particular food commodity in the food/virus combination	Grams / Kg body weight / day[22]
C6	(Degree of) Secondary spread of virus	No of secondary cases per primary case per virus

It should be emphasised that the suggested measurable units are only exemplary suggestions to demonstrate the applicability of such approach. A full risk ranking exercise would allow for a structured expert judgment elicitation to elicit the most suitable measurable units and scales.

Precise definition of each foodborne virus/berry fruits and leafy green vegetables combination in terms of the identified discrete, measurable criteria is an exercise that is usually performed with one or two experts who, ideally, would not participate in the final ranking exercise. An example of how some of such definitions could be populated for a couple of foodborne virus/food commodity combinations is shown in Table A12[23]:

---

[19] In total reported this far.

[20] For example, some information on different foodborne viruses outbreaks is compiled in a report by the Advisory Committee on the Microbiological Safety of Food - Ad hoc Group on Foodborne Viral Infections: An update on viruses in the food chain [Online - available for public consultation] Available at: http://multimedia.food.gov.uk/multimedia/pdfs/consultation/viruses-foodchain-acmsf.pdf

[21] Data is not yet available.

[22] From the EFSA Comprehensive European Food Consumption Database [Online] Available at: http://www.efsa.europa.eu/en/datexfoodcdb/datexfooddb.htm Accessed 10 07 2014

[23] Numbers introduced in the table are only exemplary for the purpose of this case study. Experts' views in these exemplary entries were based on data in (a) Advisory Committee on the Microbiological Safety of Food - Ad hoc Group on Foodborne Viral Infections: An update on viruses in the food chain [Online - available for public consultation] Available at: http://multimedia.food.gov.uk/multimedia/pdfs/consultation/viruses-foodchain-acmsf.pdf, (b) Maunula *et al*., 2013, (c) EFSA 2011 scientific opinion, (d) EFSA 2013 scientific opinion. In a full risk ranking exercise the elicitation of experts' views would be elicited with uncertainty ranges.

	C1	C2	C3[24]	C4	C5[25]	C6
Scenario 1[26]	4	46		N/K		
Scenario 2[27]	3	0		N/K		

Table A12. Example.

It is advisable that experts who participate in the ranking exercise do not see a verbal description of each scenario to avoid any biases. The verbal description of the scenarios however will be apparent to the risk analyst and the risk managers, and will be accessible to all after the exercise is completed. Also, scenarios should not be dominant (i.e., a scenario that would "score" higher than any other in all criteria).

Identification and engagement of experts for the exercise: it is anticipated that certain expertise may be needed for this case study, including (a) technical knowledge on foodborne viruses, in particular Norovirus, HAV, and HEV, (b) particulars of global trade of berry fruits and leafy green vegetables, (c) food technology, (d) relevant regulatory frameworks, and (e) risk assessment. EFSA has own protocols and procedures to identify experts who could address required expertise; a recent guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment includes further useful information on this (EFSA, 2014b).

Elicitation of experts' preferences in terms of potential risks and aggregation of results: Experts' preferences in this approach are elicited individually. These can be elicited either during structured experts' workshops or remotely. Remote elicitation of experts' rankings can be realised either via conventional excel spreadsheets or customised word templates, or via a web based facilitated tool. The latter requires a software developer to customise such tools. Individual rankings can be aggregated, and outputs can be displayed in graphs that would illustrate experts' variability in the rankings.

---

[24] This criterion could be filled in by experts much easier for HEV as incidence data do not exist for either norovirus or HAV.

[25] Information can be found in: EFSA Comprehensive European Food Consumption Database [Online] Available at: http://www.efsa.europa.eu/en/datexfoodcdb/datexfooddb.htm Accessed 10 07 2014 – but one would need to identify the food categories applicable to each scenario.

[26] Scenario 1: Norovirus in fresh/frozen berry fruit.

[27] Scenario 2: Hepatitis A in fresh/frozen berry fruit.