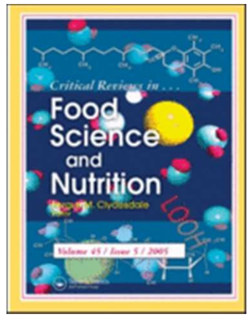


Critical review of methods for risk ranking of food related hazards, based on risks for human health

Journal:	*Critical Reviews in Food Science and Nutrition*
Manuscript ID	BFSN-2015-1828.R1
Manuscript Type:	Review
Date Submitted by the Author:	n/a
Complete List of Authors:	Van Der Fels-Klerx, H.J.; RIKILT WageningenUR, Van Asselt, ED; RIKILT WageningenUR, Raley, Marian; Newcastle University, School of Agriculture, Food and Rural Development Poulsen, Morten; Technical University of Denmark, National Food Institute Korsgaard, Helle; Technical University of Denmark, National Food Institute Bredsdorff, Lea; Technical University of Denmark, National Food Institute Nauta, Maarten; Technical University of Denmark, National Food Institute d'Agostino, Martin; Fera Science Ltd. (Fera), National Agri-Food Innovation Campus Coles, David; Newcastle University, School of Agriculture, Food and Rural Development Marvin, HJP; RIKILT WageningenUR, Frewer, Lynn; Newcastle University, School of Agriculture, Food and Rural Development
Keywords:	Risk prioritization, risk ranking, food safety, nutritional hazards, health impact

Wageningen, 8 January 2016

Dear Editor,

We would like to thank you and the reviewer for the valuable comments and suggestions to our manuscript entitled "Critical review of methods for risk ranking of food related hazards, based on risks for human health" which we submitted to Critical Reviews in Food Science and Nutrition. We appreciate a lot the suggestions given by the reviewer to improve our manuscript.

We have revised the manuscript duly taking into account each comment made. In the Annex you will find the itemized list of our revisions and responses. All co-authors have seen and agree with the revisions.

We hope you will appreciate our revisions and approve the revised manuscript for publication. In case of any question, please do not hesitate to contact me on the address indicated below.

Sincerely, Ine

HJ (Ine) van der Fels-Klerx, MSc, PhD
RIKILT Wageningen UR
PO Box 230, 6700 AE Wageningen, The Netherlands
Phone +31 317 481963
Fax +31 317 417717
E-mail: ine.vanderfels@wur.nl
Website: www.wageningenUR.nl/en/rikilt

Annex. Itemized list of responses.

Reviewer: 1

1. I urge authors to strengthen the discussion based on the findings of the literature review to provide readers with more than just an expose of the current methods available to rank risks. As it is mentioned in the manuscript, there is not a single method that can be applicable to risk ranking, but the authors must expand on this and provide directions on how to select an appropriate method for the goals of prioritization. A discussion on the differences of microbial, versus chemical and nutrition is also necessary – is there any of method that is more suitable to a certain type hazard or situation? Is it realistic (feasible) to think about a single method to rank microbial, chemical and nutrition risks? The strong discussion and conclusion are crucial and need to be included in the paper, to set it apart from the previously published report.

Answer: Yes, we agree with the reviewer to expand on the issues of how to select an appropriate method; difference of methods for microbial, versus chemical and nutrition hazards etc.
Adapted: In the revised version, we have added a strong discussion section, and wrote a stronger conclusion. To do so, we added a separate discussion & conclusion section to the paper addressing the issues mentioned by the reviewer as well as data needs of the methods; uncertainty; resource demands and communication.

2. Another concern of this reviewer is the search strategy used and the fact that it seen to have missed at least three relevant risk ranking work. The FAO/WHO produce ranking (FAO/WHO, 2008) , the US Food and Drug Administration (FDA) produce risk ranking tool, and the COI report on foodborne illness from the USDA Economic Research Services (ERS, 2015), were not included in this review, but must. The work above are not necessary different methods, but are relevant enough to be included in this review. The FDA's fresh produce risk ranking tool deserves a special attention as it is the methodology behind FDA's rule on tracking high risk foods and offers a free online tool for ranking risks in produce . It is not clear if those references were not identified at all by the search or if they were excluded from the final list of candidates. Either way, it raises the question of whether other relevant work was not excluded in this process. This review would like to receive assurance that the search strategy was robust enough to not have missed other relevant work.

Answer: If appears that the reviewer is not sure about the search strategy used in our study because three reports/papers he/she knows are not in the reference list of the paper. We would like to stress that not all references deemed relevant are given as examples in the body text and thus present in the paper's reference list.
The FDA risk ranking tool, published by Anderson et al. (2011) has certainly been included in the review, classified as a MCDA method. It was however not provided as an example to the text and thus present in the reference list. The same goes for the FAO/WHO (2008) report on produce ranking. This report has been included in our review, but was not given as example in the body text.
Adapted. In the revised version, the FDA method and the FAO/WHOI report have also been addressed in the body text. Both studies have been added in the section of the their respective method category, being MCDA and expert judgment.

The COI report from the USDA is published in the year 2015, which was out of scope of our literature study (which included publications up to and including 2013). The scientific paper (Hoffmann et al., Journal Food Protection 2012), that was published as part of the USDA study, was included in our study as relevant paper. The methodology was CoI and QALY's.

3. How each of the methods were classified is a little. For example, WTP, COI and HALY are, for this reviewer, a metric for risk ranking, not method. Authors should define better why and how they choose to classify the methods into those 14 categories, since there are many ways it could have been done.

Answer: To the opinion of the authors, a methodology is a way of doing something, in particular doing it in a systematic way, with logical steps/arrangements. Therefore, CoI, WTP and HALY were considered methods. The methods were divided into different categories based on the way they evaluated the hazards present and its severity as well as their combination to come to an assessment of the risk.

Adapted: In the revised version, this has been made more clear, by adding the following sentence "All methods covered both presence of the hazard and its severity. Method categories differed in the way in which these two factors were evaluated and combined to come to an estimate of the risk."

4. ...Authors must review the entire section on MCDA and make the necessary corrections. This reviewer recommends using as examples of MCDA methods from the papers published by Ruzante et al. (2010) and Fazil et al. (2008). Authors will see that preference functions (in addition to weights) are core to MCDA methods and must be selected when conducting a risk ranking. There are also several methods under the MCDA umbrella, which vary in complexity and might even allow for probabilistic modeling and sensitivity analysis. In addition, each of the methods has their own algorithm to calculate the "net flow," being more than just an addition (or multiplication of scores).

Adapted: In the revised paper we have rewritten the entire section on MCDA methods, such to do the corrections and to strengthen that both weights and preference functions are core part of the method, and should be selected when conducting a risk ranking. The recommended citations were included as examples to the text.

5. Line 16 and 646: this is not a systematic review, but a literature review.

Adapted

6. Line 44: the statement in this line refers to practice or is it theoretical? Please make it clear.

Adapted, we added "both in practice and from theoretical calculations" to the sentence.

7. Lines 48 to 50: include the FDA tool for produce ((http://foodrisk.org/exclusives/rrt/) and give the exact url for iRISK. Also authors should make sure they list these tools again under the method they belong.

Adapted, the section on MCDA methods has been extended to mention the FDA tool as an example of the MCDA method. The following sentence has been added: "A well-known example of a MCDA method for ranking pathogen-produce combinations is the Pathogen-Produce Pair Attribution Risk Ranking Tool (P^3ARRT) developed by FDA (Anderson et al., 2011), which is free available (http://foodrisk.org/exclusives/rrt)." Also, the URL for the iRISK tool has been corrected.

8. Line 96: was the check random? If not please state how it was done and make it clear.

Adapted. We have added "randomly selected" to the sentence.

9. Line 118: what the authors mean by type of tool? Please add between parentheses.

Adapted. The "type of tool" refers to a short description of the method or tool applied. This has now been indicated between parentheses.

10. Line 144 - 148: make sure that in the text authors follow the order stated here. This list of methods do not match the text that comes after.

Adapted. The order of the sections describing each of the method categories has been changed so to follow the order stated here. This implies that several entire sections have been moved.

11. Line 198: make sure the subheadings are consistent throughout the text – see line 198 and 234, for example. And on this particular title for the subheading, it is really focused on the risk manager, not on the broad group of stakeholders.

Adapted. Subheadings have been made consistent, and focused on the risk managers. So, we used "Perspective for use by risk manager" as subheading.

12. Line 204: please make it more clear what this method entails. It was extremely confusing to this author how it differs from just risk assessment. In my field of work, for example, comparative risk assessments are the same as relative risk assessments (see lines 178-179), but according to your review, CRA is a different method that seems to restrict the comparison to fatalities. Please clarify the distinction between risk assessment and CRA.

Answer: In our study, comparative risk assessment were defined as methods that use population attributable factors to estimate total effects of a risk factor – in this case a food related hazards on numbers of dying related to diseases caused by that risk factor. CRA make use of large epidemiological dataset. They clearly distinct from RA and relative RA since they are not based on the total consumption of the hazard (via food). The term 'comparative' could indeed by used in different ways in literature, in this case it is not identical to 'relative'. Indeed, the part on relative risk assessment was missing in the original paper, though covered in the introduction.

Adapted: We have one line to the CRA section to clarify the focus of CRA in our study: " CRA is restricted to comparisons of deaths and it is, therefore, not comparable to a risk assessment or a relative risk assessment." *Also, we have moved the lines on relative risk assessment from the introductions, to the section on RA.*

13. Line 239: please mention whether this is a qualitative, semi-quantitative or quantitative method.

Adapted. the sentence has been changed into : "Risk ratios or quotients refer to a quantitative method in which estimates of exposure are divided by estimates of effect".

14. Line 263: lack of data seem to be an issue for all methods. If some are better than other in dealing with this, please make the distinction, otherwise it worth mentioned up front instead of under each of the methods.

Yes, the reviewer is correct. Lack of data seems to be an issue for all methods. However, for some methods it is more an issue than for others, particularly for RA and CRA and MCDA. In the section referred to by the reviewer, it is not so much an issue of the three methods mentioned and, therefore, we have deleted the two sentences on lack of data here. In the discussion, we have added an entire section on the data needed by the different method categories, and if they can deal with lack of data.

15. Line 296: typo – should be "and".

Adapted

16. Line 349: instead of "may be advisable" should say "is advisable".

Adapted.

17. Line 349 -350: updating ranks as new information becomes available is also a general issue with all methods. As for the comment above, this is not the case for some of the methods, please note otherwise stick to a general weakness statement in the beginning or end of the article.

Adapted. The statement of updating ranks as new information becomes available has been removed from the COI section. Instead, it has been placed in the general discussion section, but referring as a strength of all methods to which this is applicable. As part of the new discussion section, the following sentence has been added " Methods most suitable for such an automatic update are RA, risk ratio, risk scoring, risk matrices, COI, HALY, and MCDA. It is more difficult to apply with CRA, WTP and expert synthesis".

18. Line 376: Newsome et al (2009) and Chen et al (2013) are the same method – iRISK. Use just one.

Adapted.

19. Line 378: one of the issues of DALY or QALY is also communication – it is hard for stakeholder to understand that they mean – please list that as a weakness too.

Adapted. The following sentence has been added "Also, stakeholders have difficulty to understand the concept and what is meant by it".

20. Line 483: Havelaar et al. (2010) is not on the reference list – this reviewer did not check all the references, but please make sure they are all there.

Adapted. Havelaar et al (2010) has been added to the reference list. Also, all other references have been checked and added/corrected.

21. Line 521-522: are those subjective? Please make it clear how risk classes are established in this method.

Adapted. Yes indeed, those are subjective. This has been made clear by adding the sentence "The division into these classes is subjective." Furthermore, we added the following line in the paragraph on strengths and weaknesses of this method. "However, the division between different categories for presence of the hazard (e.g. low, medium high occurrence) and its effects (e.g. low, medium, high toxicity) is subjective and, thus, other results are obtained when with other divisions."

22. Line 531: an extra "I" before "Alternatively".

Corrected.

23. Line 536: experts to do what? Please finish the sentence

Adapted. Sentence is confusing and therefore removed.

24. Line 595-596: in MCDA judgement of stakeholders are not used to rank risks directly, but are inputs on how to weight the different criteria and in establishing the preferences.

Adapted. This has been added when rewriting the MCDA section.

25. Line 600: FAO/WHO produce risk ranking must be mentioned here too.

Adapted. The FAO/WHO produce risk ranking method is presented as an example in the section on expert synthesis.

26. Line 651-652: what are those methods that allow for microbial and chemical to be ranked together? List here.

Adapted. In the revised paper, the discussion section is extended. The following line has been added to the discussion section: " Four of the eleven method groups can be applied to all three types of hazards (microbiological, chemical and nutritional), either alone or in combination, being MCDA, risk matrices, stated preferences, and expert synthesis."

27. Line 658: MCDA are extremely data intense (see Ruzante et al., 2010 and Fazil et al., 2008) – it all depend on your criteria.

Adapted. We agree MCDA are data intense, and have removed MCDA here.

28. Line 644: need to the stressed in the conclusion that uncertainties need to be clearly stated as the majority of those methods do not provide this strength.

Adapted. A sentence has been added to the conclusion stressing the importance on clearly stating the uncertainties in data input.

29. Table 3: this author disagree that MCDA methods require a moderate amount of resources. Establishing weights and preferences with decision makers and getting the necessary data to run the analysis is extremely time consuming. MCDA can be a quite robust quantitative method, with even stochastic version – the authors seem to have a very simplistic view of what MCDA method is. Graphs are another method for communication for MCDA methods. And for COI, HALY and MCDA, the data needs expressed on the last five rows of the table would be correct if the approach been taken is "top-down," but incorrect if using "bottom-up", in this case you would need all of the information mentioned in the last rows (see who iRISK works).

Adapted. We agree with the reviewer that MCDA requires a high amount of time, data and money, and have adapted this in Table 3. Also, graphs have been added as a method of communication for MCDA methods.
Table 3 provides essential data needs. This has been changed in the heading. Indeed CoI, HALY and MCDA, can also use some of the other data sources mentioned when the essential data is missing, and thus taking the bottom-up approach but this is less efficient.

RESEARCH PAPER

Critical review of methods for risk ranking of food related hazards, based on risks for human health

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], M. D'Agostino[4], D. Coles[2], H. J. P. Marvin[1] L. J. 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; [3]Technical University of Denmark, National Food Institute, Morkhoj Bygade 19, 2860 Soborg, Denmark; [4]Food and Environmental Research Agency, Sand Hutton, York, North Yorkshire, YO41, 1 LZ.

ABSTRACT

This study aimed to critically review methods for ranking risks related to food safety and dietary hazards on the basis of their anticipated human health impacts. A literature review was performed to identify and characterize methods for risk ranking from the fields of food, environmental science and socio-economic sciences. The review used a predefined search protocol, and covered the bibliographic databases Scopus, CAB Abstracts, Web of Sciences, and PubMed over the period 1993-2013.

All references deemed relevant, on the basis of of predefined evaluation criteria, were included in the review, and the risk ranking method characterized. The methods were then clustered – based on their characteristics - into eleven method categories. These categories included: risk assessment, comparative risk assessment, risk ratio method, scoring method, cost of illness, health adjusted life years, multi-criteria decision analysis, risk matrix, flow charts/decision trees, stated preference techniques and expert synthesis. Method categories were described by their characteristics, weaknesses and strengths, data resources, and fields of applications.

It was concluded there is no single best method for risk ranking. The method to be used should be selected on the basis of risk manager/assessor requirements, data availability, and the characteristics of the method. Recommendations for future use and application are provided.

KEY-WORDS

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

1. INTRODUCTION

Ranking of health risks related to food safety and nutrition is generally recognised as the basis for risk-based priority setting and resource allocation. It permits governmental and regulatory organisations to allocate their resources efficiently to the most significant public health problems (Van Kreijl et al., 2006). Within the area of food, risk is defined as the analysis and prioritization of the combined probability of food contamination, consumer exposure and the size of the anticipated public health impact of specific chemical, microbiological and/or nutritional hazards related to food. It is the combination of the *probability that a hazard may occur* in a food product and the *effect of exposure to the hazard on human health* (Codex Alimentarius 2001). Risk ranking has been applied to food safety monitoring programs and has shown to increase the efficiency of monitoring and to decrease inspection costs, both in practice and from theoretical calculations (Baptista et al., 2012; Presi et al., 2008; Reist et al., 2012).

To date, various risk ranking methods are available that prioritise food safety risks (Van Asselt et al., 2012). Methods vary from qualitative, through semi-quantitative, to quantitative methods (Cope et al., 2010; Van Asselt et al., 2012). Most methods are based on the 'technical' concept of risk being a function of presence of the hazard and severity of its impact on human health. However, some methods also involve other metrics, which may be considered in decision making, e.g., consumer perceptions of risk. In order to determine which methods are most suitable for ranking food related risks, it is important to follow a structured, objective and transparent approach to identifying and evaluating the available methods (van Asselt et al., 2013).

The aim of the current study was to review available methods for ranking risks associated with food on the basis of anticipated health impact, to characterize the methods and to provide recommendations for their use.

2. MATERIAL AND METHODS

2.1 Protocol for literature review

A literature review was conducted which aimed to identify risk ranking methodologies that can be used to prioritize food related hazards, on the basis of the size of anticipated health impact. Hazards are defined as those agents that can be present in food and can negatively affect human health (Codex Alimentarius, 2001). Hazards included in this study were nutritional, chemical and microbiological hazards. The review covered methods from the fields of natural/life (food) science, socio-economic sciences and food safety governance, published during the period 1993-2013. Risk ranking methods from fields outside food science (i.e. environmental sciences and socio-economic methods) were also included to evaluate their appropriateness for application in food science. The literature review followed the principles of a systematic literature review as described by EFSA (2010). A protocol for the structured literature review was defined *a priori*, including search strings and criteria for evaluation of the literature references (Annex 1).

2.2 Literature review

Review methodology

a. Scientific articles were identified using the following bibliographic databases: Web of Science, Scopus, PubMed, and CAB Abstracts. In addition, the general search engine Google was used to search for reports, (the 'grey literature'), from relevant international and national organisations, authorities, and agencies (e.g., EFSA, EMA, WHO/FAO, FDA, Health Canada, OECD). The literature search focused on papers and reports published in English.
b. The set of search strings was applied leading to an initial set of search results. All retrieved references were stored in an Endnote database. Duplicates, a result of using four different bibliographic databases, were removed.
c. The references resulting from the initial set of search results were screened for their relevance to the study objectives by applying the 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 allocated to one of three categories and placed in the corresponding category of the Endnote database:

- *Relevant for this study*: the reference was included;
- *Possibly relevant for this study*: uncertain if the reference was relevant for the study;
- *Not relevant for this study*: the reference was determined to be out of scope.

An inter-observer check was conducted with a randomly selected subset (10%) of both selected and excluded references.

d. In tier 2, the full text of the references that were in the *Relevant* and *Possibly relevant* groups of the Endnote database were retrieved. By reading the full texts, the papers/reports were evaluated for their relevance to the field of interest and their quality using the evaluation criteria. When deemed relevant, the reference was retained or moved to the group *Relevant* in the Endnote database. When deemed not relevant, the reference was moved to the group *Not relevant* in the Endnote database. Also at this stage, an inter-observer check was conducted; certain (randomly chosen) literature references were evaluated by two experts from the team (from different disciplines) in order to gain insights into the variation between the evaluation results of two different experts.

e. Citations used in the reports/references of the final Endnote database were screened for additional relevant references, published after 1993 (snowball citation), and steps c) and d) were applied to them.

Evaluation of references

For each reference stored in the *Relevant* category of the Endnote database, the risk ranking method and its characteristics were evaluated in depth. A summary of the information obtained was stored in an excel sheet, using a unique row for each reference. The format of the excel sheet was defined beforehand, starting from the template developed by EFSA's BIOHAZ panel (EFSA, 2012b), but with some modification to increase relevance to the objectives of the current study. Separate columns were utilised for information about the reference (author names, title, abstract, journal, volume and page numbers), and for storing the results from the critical evaluation of the risk ranking methods including: the type of tool (short description); field of application (microbiological, chemical, and/or nutritional hazards); 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 requirements that describe the model variables (e.g., human population data, or microbial numbers); method of data collection, describing how the necessary data were collected and which data sources were used, and finally data integration, describing how data were integrated in the application described in the reference. Based on this evaluation, the references and the evaluated methods were categorised into different groups of methods. The method categories were then described according to the following characteristics: scope, application area, approach, strengths and weaknesses, and perspective for use by risk managers. At this stage, reviews on risk ranking methods and other relevant literature were also consulted..

3. RESULTS

3.1 Literature search

At tier 1, application of the search strings and removal of duplicates led to the retrieval of the following numbers of references (Table 1): 6021 for chemical/toxicological hazards; 2932 for microbiological hazards; 1049 for nutritional hazards; 112 references using health adjusted live years method; and 3358 references using socio-economic methodology. The latter two method groups were considered since they could potentially include each of the three types of hazards (microbiological, chemical and/or nutritional hazards). The total numbers of references appearing in tier 2 are somewhat higher than in tier 1 due to snowballing citations. In total 253 references were judged to be relevant.

3.2 Description of risk ranking methods

Based on the evaluation of the methods described in the relevant references, the risk ranking methods were classified, according to methodology, into the following categories: 1) Risk Assessment (RA), 2) Comparative risk assessment (CRA), 3) Risk ratio method, 4) Scoring method, 5) Risk matrix, 6) Flow charts (including decision trees and influence diagrams), 7) Cost of illness (CoI), 8) Health adjusted life years (HALY), 9) Multi criteria decision analysis (MCDA), 10) Stated preference methods, and

11) Expert judgement. Table 2 shows the numbers of references that presented a particular method category, per type of hazard. All methods included both presence of the hazard and its severity. Method categories differed in the way in which these two factors were evaluated and combined to come to an estimate of the risk. In some instances, a combination of methods was applied, in which case the study was classified to its main category.

RA was by far the most frequently applied method. This method was applied to both chemical and microbiological hazards. For each of the chemical and microbiological hazards, about one third of all tier 1 references described the application of a RA to a particular hazard. However, as the procedure for each of the chemical and microbiological RA is comparable, only references describing guidelines for performing a RA were included. Risk ratio, scoring, risk matrices and flow charts were mostly applied to chemical hazards, whereas CoI, HALY, and expert judgments were mostly used for ranking microbiological hazards (Table 2). Ranking methods for nutritional hazards were fewer, and were mostly based on RA, CRA and expert judgement (Table 2). CRA, CoI, and stated preferences were the methods that were applied least frequently, with CRA used in three studies about nutritional hazards, and the latter two methods primarily applied to microbiological hazards. A few studies have considered both chemical and microbiological hazards in their ranking, applying methods for CoI and HALY. Summaries of each method and characteristics are presented in the following sections and in Table 3.

<u>3.2.1. Risk Assessment</u>

<u>Scope:</u> A RA for a chemical or microbiological hazard aims to estimate the risk for human health associated with the presence of the hazard in one or more food products, and total food consumption. Numerous risk assessments have been applied to chemical and microbiological hazards in food. WHO (WHO, 2009) and Codex Alimentarius (2014) have provided guidelines regarding the principles and methods for the risk assessment of chemical contaminants and pathogens in foods. Although the application of the RA methodology is tailored to the hazard type, the principles for performing a risk assessment for both types of hazards are identical, consisting of the following four steps: hazard identification, exposure assessment, hazard characterisation, and risk characterization.

Application area: Risk assessment is usually applied for one identified (chemical or microbiological) hazard occurring in a specific food commodity and for a predefined population, with the purpose of characterizing the associated health risk. Apart from this, an important reason for conducting a RA is to evaluate the impact of control measures to reduce the risk. If the results of different RA are compared (e.g. for different hazards or different foods), the RA can be used for risk ranking.

Approach: Various RA approaches for chemical and microbiological hazards in food were identified, applying different combinations of deterministic, probabilistic (or stochastic), qualitative, semi-quantitative, and quantitative modelling. Furthermore, different approaches were used for the exposure assessment and the hazard characterization steps. EFSA (2011) published an overview of procedures for current RA methods for dietary exposure of different chemical substances. The need for development of harmonized approaches, and future exploration of cumulative exposure assessments, is identified. In 2012, EFSA published its experiences gained with Quantitative Microbiological Risk Assessment (QMRA) studies (EFSA, 2012a).

Strengths and weaknesses: In RA, all available scientific and technical information and data, as well as variability and uncertainties are systematically organized and analysed. It is a well-structured method, providing insights into what is known and what is not known. In particular, RA offers the opportunity to address uncertainties in a transparent way, e.g., *via* sensitivity analyses and/or modelling and simulation runs. It could be the most precise method to estimate risks, including the relevant uncertainties. However, a RA for one chemical or microbiological hazard usually requires a lot of time, data and knowledge. Ranking risks related to various hazards in food using outcomes of individual RAs will take even more resources and RAs are often hampered by a lack of quantitative data. Lack of data, selection of models to fit to the data, and assumptions that need to be made give rise to uncertainties in the outcomes. Recently, several tools for relative risk assessment for pathogens of pathogen-food combinations have been published. Examples of such tools applying quantitative methods are the swift QMRA tool (Evers and Chardon, 2010) and iRISK, which is a relative risk assessment system for evaluating and ranking food-hazard pairs (Chen et al. 2013, see http:// https://irisk.foodrisk.org/). An example of a semi-quantitative approach is Risk Ranger (Ross and Sumner, 2002) developed by Food Safety Centre (2010).

Perspective for use by risk manager: Applied optimally, RA should disseminate key information regarding risk from exposure to food hazards to policy makers, decision makers and the public. RA are very useful for providing insights into gaps in knowledge and issues associated with high levels of uncertainty. However, they may not be suitable for risk ranking given the large amounts of data, knowledge and resources needed.

3.2.2. Comparative risk assessment

Scope: A Comparative Risk Assessment (CRA) analysis can estimate the number of deaths that would be prevented in a given period if current distributions of risk factor exposure were changed to a hypothetical alternative distribution (Danaei et al., 2009; Micha et al., 2012). In these papers, CRA is restricted to comparisons of deaths and it is, therefore, not comparable to a risk assessment or a relative risk assessment.

Application area: Three applications of CRA have been found; each of them studied the impact of dietary factors on disease mortality. 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. Lim and co-workers (2012) investigated burden of disease and injury attributable to 67 risk factors (including chemical hazards and nutritional imbalances) in 21 regions through application of a systematic analysis for the Global Burden of Disease Study 2010. Although a CRA analysis as described below was not performed by Lim et al. (2012), several elements of a CRA analysis were included.

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. Input needed to determine the PAF include: a) effect size (relative risk estimate) of the causal diet-disease relationship, b) optimal or theoretical minimum-risk exposure distribution, c) dietary risk factor exposure distribution in the population and, d) total number of disease-specific deaths (plus non-fatal events, when available) in the population. Data sources for obtaining these inputs include epidemiological studies, systematic reviews, meta-analysis, nationally representative nutrition surveys and mortality databases.

Strengths and weaknesses: A CRA analysis is a systematic assessment of unbiased data collected in national and international surveys as well as the peer reviewed literature. It allows for consistent, comparable and quantitative assessment of the global impact of risk factors on disease by sex- and age-specific groups. A CRA analysis requires knowledge and resources (manpower, money, data), which makes it expensive to perform. Unbiased data are also needed, e.g., to establish exposure distributions or causal diet-disease relationships, which may often not be easily accessible or available. The weights of different diseases are not considered. Uncertainties associated with a CRA analysis can be high because of data limitations.

Perspectives for use by risk manager: A CRA analysis offers a global assessment of the impact of dietary factors on disease mortality, which is very valuable for priority setting and policy making. However, with large and overlapping uncertainty ranges for the different risk factors, ranking of modifiable dietary risk factors may be difficult.

3.2.3. Risk ratio method

Scope: Risk ratios or quotients refer to a quantitative method in which estimates of exposure are divided by estimates of effect. For this purpose, data are needed regarding the amounts of the hazard consumed (either the dose or the concentration) as well as a measure for the effect of the hazards that are studied.

Application: The risk ratio method has usually been applied to rapidly screen the risk of a range of chemical compounds in order to rank them. Most studies applied the method to rank pesticides, although five studies focused on microbiological hazards, and one study applied the method to rank both chemical and microbiological hazards.

Approach: For chemical contaminants, 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) (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 (Bench Mark 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 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). Applications of the method to microbiological hazards used different criteria, such as costs and effective dose.

Strengths and weaknesses: This method is easy to understand, and can be applied once concentration data and toxicological reference values are available; it only needs an estimate for both amounts of the hazardous material consumed and the effect of the hazard on human health. For emerging chemical hazards, e.g., nanomaterials, toxicological reference values are usually not available. .

Perspectives for use by risk manager: The method can give a quick answer on the risk of food safety hazards for human health, and can be applied to both chemical and microbiological hazards.

3.2.4. Scoring method

Scope: This method is based on semi-quantitative scoring of both exposure and effect of the hazard on human health, followed by their multiplication (or – in one reference - addition).

Application: Scoring methods provide a simple risk ranking method to 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).

Approach: When a scoring method is applied, both exposure and severity (or effect) endpoints are considered. However, endpoints for exposure and effect can vary. Various endpoints have been used to estimate exposure, such as chemical transformation properties (degradability, half-life), mobility/distribution (such as bioaccumulation factors (BAF) or bioconcentration factors (BCF)), release, 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, selection of appropriate endpoints for a specific study is one of the steps in ranking risks according to a scoring method. Examples of endpoints for effect on human health might include acute toxicity, carcinogenicity, or reproductive toxicity, and can be based on LD50, MOAEL, BMDL10 etc. Once criteria are set, endpoints are classified semi-quantitatively, e.g., using scores from 1 to 3 or from 1 to 5, as applied in, for example Penrose et al. (1994).

After this classification system for endpoints has been established, data sources need to be found in order to assign scores for exposure and effect. These sources can be based on literature, available data and/or expert opinion. Scores subsequently need to be aggregated, which is mainly done by multiplying exposure and effect (see, e.g., 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 employ 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., 2005). 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.

Strengths and weaknesses: This semi-quantitative method is easy to conduct once scores have been assigned to the model variables. Furthermore, it allows the inclusion of stakeholder perceptions in assigning the scorings and the importance (to each stakeholder) of each model variable is reflected by the weighting allocated to it. The assigned weights should then be clearly documented to guarantee a transparent approach.

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

3.2.5. Risk matrices

Scope: Just like the scoring methods, risk matrices also make use of scoring both exposure and effect endpoints. The difference between scoring methods and risk matrices is that, in the latter, the exposure and effect endpoints are not aggregated by multiplication or addition, but are 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 or microbiological hazards for which limited quantitative data are available. This method has, for example, been applied for ranking the risks of nanomaterials (O'Brien and Cummins, 2011; Sorensen et al., 2010; Zalk et al., 2009).

Approach: Both the likelihood of occurrence and the consequences of the hazard for human health are scored into one of several classes; see Figure 2 for an example. Classes that could be used for likelihood of occurrence are: almost certain, likely, possible, unlikely and rare. Classes that could be used for the consequences are: insignificant, minor, moderate, major and severe. The division into these classes is subjective. Then, risk classes are assigned to the combinations of Likelihood and Consequences, e.g., being L (low), M (moderate), H (high), and E (extreme), as shown in Figure 2. Risk classification may also be based on scores. Zalk et al. (2009), for example, classified nanomaterials based on scores for probability and severity, and the results were depicted in a risk matrix. The results can also be visualized using spider web plots, as conducted by, (e.g)., Ranke and Jastorff (2000), who classified various endpoints using scores from 1-4, and compared plots for the various compounds to obtain an indication of the most risky ones.

Strengths and weaknesses: The risk matrix method is qualitative or semi-quantitative, and thus less accurate than methods based on concentration data and dose-response relationships or toxicological reference values. It provides a visualisation for both presence of the hazard and its effects, giving direct insights into the way these two elements contribute to the overall risk of a hazard. For example, a hazard may present a high risk due to a high exposure, although its severity is low. Alternatively, due to its high toxicity, it may present a high risk rank despite low exposure. Matrices will give more information to the risk manager compared to other methods that produce a list of hazards according to the overall risk alone. However, the division between different categories for presence of the hazard (e.g. low, medium high occurrence) and its effects (e.g. low, medium, high toxicity) is subjective and, thus, other results are obtained when with other divisions.

Perspectives for use by risk manager: 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.

3.2.6. Flow charts

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

Application: Flow charts or decision trees can be used for various purposes. In general these methods are used to obtain a qualitative indication about the risks associated with hazards. Haase et al. (2012), for example, established a decision tree for nanoparticles to determine whether a full risk assessment is required or not. EFSA described guidelines for classifying chemical hazards as negligible, low, medium, and high risks (EFSA, 2012c, 2012d).

Approach: A flow chart is generally based on several questions that need to be answered in order to arrive at a certain risk class. Questions can be based on the likelihood that specific chemicals or microbiological hazards are present in the study object; evidence of occurrence or incorrect practice in the food chain, the toxicological profile, and the outcome of national monitoring programmes (EFSA, 2012c, 2012d). Eisenberg and McKone (1998) used a Classification and Regression Tree Algorithm (CART) to specify the chemical and environmental properties and Monte Carlo simulations to estimate human exposure. Schmidt et al. (2011) utilized a decision support system (DSS) to rank genetically modified organisms (GMOs), based on a decision tree and rules, indicators and baselines, and thresholds (such as the LD50) (Schmidt et al., 2011). DSS may also be combined with multi-criteria decision analysis (MCDA). Critto (2007), for example, utilised a DSS system to evaluate ecological observations and ecotoxicological tests for contaminated sites and then incorporated MCDA and expert judgments into the ranking. This approach might also be used for ranking food safety risks.

Strengths and weaknesses: Flow charts/decision trees present a straightforward method with clear questions for which only qualitative information is needed, although quantitative information can be used where available. The method can, thus, be used for a quick screening of food safety hazards, in order that the most relevant ones may subsequently be investigated in more detail. However, this method strongly depends on expert input and it is, therefore, essential to perform a rigorous expert elicitation study. Furthermore, this type of method is vulnerable to being less transparent than other methods, as it is not always clear why hazards end up being classified as a high, medium or low risk. Therefore, for each hazard 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 risk manager: It is important to set up the right questions for inclusion in a flow chart/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.

3.2.7. Cost of Illness method

Scope: The underlying research objective of the Cost of Illness (CoI) approach is distinct from those of the methodologies described so far. CoI studies acquire data for conducting economic analysis in order to obtain a ranking in terms of how society might allocates scarce resources when addressing food-related hazards. The procedure involves calculating the directs costs to society related to disease and death due to chemical, microbial and/or nutritional 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 loss of income. Subject to data availability, it is possible to compare large numbers of food risks.

Application area: 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).

Approach: 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 experiences each type of impact, defined as the disease severity class. A critical point is whether it is restricted to acute effects, or whether long-term effects (sequelae and deaths) 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).

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 (e.g., Batz et al., 2011) or frequency distributions can be used in Monte Carlo or stochastic simulation models (Lake et al., 2010; Kemmeren et al., 2006). The costs incurred at each state are calculated, often including the categories of direct health costs, indirect health costs, and indirect non-health costs.

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 one year, whereas other end-points, such as life-long disabilities, are strongly affected by discounting. Hence, the effect of discounting will differ per hazard (Kemmeren et al., 2006) and the rate of interest selected.

<u>Strengths and weaknesses</u>: The CoI method employs readily available and reliable data (Buzby et al., 1996) and the calculations are transparent and relatively simple. The same disease incidence data are used in HALY calculations so it is relatively efficient to produce both sets of rankings at the same time and they are, to some extent, complementary. A combined risk ranking can also be produced. 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, the measures excludes non-workers, and do not address perceived quality of life including factors such as pain and stress (Golan et al., 2005). 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. The results are dependent on the assumptions made *inter alia* about medical outcomes and the prevailing labour market.

<u>Perspectives for use by risk manager</u>: CoI is a well-tried technique with well-understood limitations relating to missing data, and failure of the approach 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 is advisable. Due to non-standardisation of technique (e.g. different components, and assumptions), comparability between studies is awkward.

3.2.8. Health adjusted life years (Burden of Disease)

Scope: 'Health adjusted life years (HALY)' 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 is then multiplied by the duration of that health state. A descriptive summary of the various HALYs is presented by Mangen et al. (2014).

Application area: HALY measures 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, i.e., the burden of disease. HALY estimates such as disability adjusted life years (DALYs) or quality-adjusted life years (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. 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 HALY approach has been applied to rank different pathogens and chemical contaminants in the same food category, 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: Data are required for estimating the number of cases with the most relevant types of acute illnesses, chronic sequelae and mortality (also termed health outcomes) arising from exposure to the hazards under consideration. Different types of hazards (chemical, microbiological or nutritional) require different types of data and modelling approaches (Crettaz et al., 2002; Hofstetter, 2002; Mangen et al., 2010; Mangen et al., 2014; Pennington et al., 2002), but after the final DALY/QALY calculations have been made, the risks estimates should be readily comparable. DALY/QALY estimates may also be included in several of the other risk ranking methods such as RA (Howard et al.

(2007); Newsome et al. (2009)), CRA (Lim et al. (2012)), MCDA (Ruzante et al. (2010)), risk matrixes, flow charts/decision trees or in expert syntheses.

Strengths and weaknesses: HALY methodologies readily allow 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 et al., 2012) and ECDCs initiative for developing methodologies for measuring current and future burden of communicable diseases (Mangen et al., 2014).

DALYs and QALYs are semi-quantitative estimates based on disability scoring, and their accuracy is highly dependent on the quality of input data and risk assessment models used for estimating the incidences of relevant health outcomes. In the applied studies, the methods for estimating the incidences of relevant health outcomes varied widely. The estimated DALY or QALY values seem to be relatively precise quantitative estimates, and there is a risk of over-interpretation of the relative differences, if the level of uncertainty is not addressed. A general methodological weakness is inadequate evidence to estimate the incidences of chronic disability, especially in cases with few or no symptoms during the acute phase of a disease. Another methodological weakness is that the concept of DALYs assumes a continuum from good health to disease, disability, and death which is independent of time – a concept not universally accepted. Also, stakeholders have difficulty to understand the concept and what is meant by it.

Perspectives for use by risk manager: Tools are readily available for calculating DALYs for a range of infectious diseases including foodborne zoonoses in the EU (BCoDE tool from ECDC). If RA or models for estimation of reported cases are available, the resources needed to estimate DALYs are moderate. However, development of RA 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 of the 'societal production losses' related to the included hazards. If HALY estimates from different studies are to be used in risk ranking, then differences in the methodology employed and the comparability of the studies must be considered. For monitoring purposes, risk ranking models estimating HALYs 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).

3.2.9. Multi-Criteria Decision Analysis (MCDA)

Scope: MCDA is an approach which has the potential to evaluate multiple - often conflicting - criteria in decision making. It allows for comparison of different risks on common basis, by simultaneous consideration of technical information, uncertainty and different stakeholder preferences, both quantitative and qualitative data, and the integration of large amounts of complex information. . MCDA helps structuring and solving problems, such to enable making more informed and better decisions. In the context of risk ranking, important criteria utilized in food safety can be identified through a process of expert or lay consultation, which may include not only public health impacts but also perception, costs – an in case of interventions – also weight of evidence, and practicality associated with the interventionsApplication area: MCDA can be applied to any range of problems, which can be defined in terms of a common set of criteria. As the scientifically 'best' solution may be inadequate in terms of acceptability to society, utilize resources which or not available, or be sub-optimal in terms of allocating resources, stakeholder methods are sometimes used to capture the preferences of consumers, citizens and/or experts. MCDA which combines expert judgement across a range of relevant criteria appears to be the second most popular method for relative risk ranking of microbiological hazards, after RA.

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, stakeholders or lay people (Fazil et al., 2008), can be supplied with technical information in relation to each risk criterion to assist their deliberations. The selection of preference functions and weights are an integral and core part of the MCDA methodology and must be selected when conducting a risk ranking. An example is provided by Ruzante et al. (2010) who utilized the method to develop a prioritization framework for foodborne risks that considered not only public health impacts but also market impact, consumer risk acceptance and perception, and social sensitivity. Another well-known example of a MCDA method for ranking pathogen-produce combinations is the Pathogen-Produce Pair Attribution Risk Ranking Tool (P^3ARRT) developed by FDA (Anderson et al., 2011), which is available free (http://foodrisk.org/exclusives/rrt). Fazil et al. (2008) applied MCDA for the ranking of food safety interventions, considering amongst others cost, effectiveness, and weight of evidence. MCDA methods and applications vary in their complexity; they may even allow for probabilistic modelling and sensitivity analyses. Recently, alternative methods for performing a MCDA have been developed and employed, e.g., by Havelaar et al. (2010), in order to minimise the biases linked with experts' direct weighting of the MCDA criteria.

Strengths and weaknesses: MCDA allows consideration of stakeholder perceptions by using the weights and preference functions they assign to the various criteria in the analysis. Furthermore, economic impact or other criteria that are deemed relevant can be included, in addition to human health criteria. 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, therefore, allows inclusion of subjective elements that may also be important for risk managers to include in their decision making processes, depending on the aim of the ranking exercise. Alternative scenarios using weights and preference functions for various input factors can be compared. However, MCDA outcomes are more difficult to communicate compared to more straightforward methods such as risk matrices or scoring methods, as various criteria are included, which are weighted and prioritized differently. Furthermore, this method needs expert or stakeholder input in order to derive the weights and preference functions for the criteria. Therefore this method has weaknesses that are linked to the elicitation of information from experts (see below), i.e., the 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); the need to consider how to elicit experts' own uncertainties regarding their views, opinions, judgments; and - last but not least – the need to consider possible ways to combine individual opinions without masking variability in the experts' views.

Perspectives for use by risk manager: This systematic method is very valuable in cases where stakeholder perceptions are required to be included in the risk ranking, as weights and preference functions can be assigned to the various model variables. This method also allows the inclusion of factors other than effect and exposure endpoints, e.g. from the social-economic field, or in terms of policy development, which makes it a very versatile tool. The application of MCDA will provide a single number for ranking. However, the underlying calculations can be difficult for the non-expert to understand for those without expertise in the methodology.

3.2.10. 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 for risk reduction. 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.

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 safety risks is limited and largely confined to valuing individual disease reduction measures or comparing alternative risk management options within single food-disease problem, see e.g., Mørkbak & Nordström (2009) and Miller et al. (2005). Golan et al (2005) concluded that, at present,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: Using stated preference methods, 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. In contrast to the CoI approach, stated preference methods include the value individuals place on other factors for which no markets exist such as, for instance, (not) experiencing pain. 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.

One of the stated preference methods, willingness to pay (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). As an example, Mørkbak and Nordström (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. This approach defines the choices which individuals make in terms of the levels of key attributes (such as high/low price, probability of illness etc) which are associated with each of the goods being compared.

Strengths and weaknesses: 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 revealing which societal groups (e.g., by age, background) ranks a particular risk most highly (see Haninger and Hammitt (2011) for an example). The aggregated value of benefits (or societal WTP) of food safety (e.g., reduced risks) can be compared with the costs for achieving them since both costs and benefits are expressed in monetary units.

However, WTP is a difficult technique to apply, and is prone to errors and bias unless conducted meticulously. Experience so far has been in comparing only 2 to 4 alternative risks. 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. Moreover, 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.

Perspectives for use by risk manager. These techniques provide a means to incorporate societal preferences in ranking and decision making. However, experience in the food safety field as yet is only modest, and there is scope to develop techniques still further.

3.2.11. Expert judgement

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.

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). The inclusion of public perceptions, priorities and values may result in a different ranking being reached to that derived from using scientific experts alone. This 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 (DeKay et al., 2005).

Approaches: A variety of methods is available, for application in workshops or in surveys, which may be characterised by the flows of information which take place between the participants and the research team (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, e.g., MCDA and the Carnegie-Mellon approach. In MCDA , the judgement of stakeholders is used to allocate weights and potentially also on the way to weight the different criteria and in establishing the preferences to the different attributes whereas the Carnegie-Mellon approach produces risk rankings. . Approaches also 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. Four approaches were identified:

- 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). Commonly used approaches are workshops and the Classical Delphi method (Van der Fels-Klerx et al., 2002).
- Survey based on existing knowledge of lay or expert participants (i.e. minimal technical communication during the study), as applied by, e.g., Schwarzinger et al. (2010) and Harrington (1994).
- 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. FAO/WHO, 2008), 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). 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).
- 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. 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 (DeKay et al., 2005).

<u>Strengths and weaknesses</u>: Judgement-based methods provide additional information to that of technical assessments, e.g., when 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. However, judgement-based methods require a 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 ned for meticulous preparation the method is 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 risk manager: 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. Also, they could provide a means of incorporating societal values into risk ranking.

DISCUSSION AND CONCLUSIONS

A literature review has been performed on methodologies for ranking risks related to chemical, microbiological and nutritional hazards in food, on the basis of their anticipated effects on human health. The results showed that a range of risk ranking methodologies has been applied depending on the purpose of the specific study. They have been grouped into eleven main categories, determined primarily by the type(s) of hazard that can be ranked, data needs, and uncertainty. Some methods allow ranking of different hazards types (chemical, microbiological), whereas others allow ranking only within one hazard category.

Four of the eleven method groups can be applied to all three types of hazards (microbiological, chemical and nutritional), either alone or in combination, these being MCDA, risk matrices, stated preferences techniques, and expert synthesis. For microbiological hazards, there is a close relationship between exposure and resulting levels of illness and death, which allows CoI and DALY/HALY calculations to be made. With chemical contamination of food, there is no such direct relationship between the contamination and resulting diseases/deaths in the population, since effects on human health are long-term and, hence, the cause-effect relationship is difficult to establish. Consequently, these methods are not often applied to chemical food contamination, although an exception is the study by Kemmeren et al. (2006) who calculated DALYs for chemical contaminants, using assumptions on the relations between chemical food contamination and disease outcomes. Although health effects of nutritional hazards are often evident only in the longer term, recent improved availability of insights from long-term epidemiological studies on the cause-relationships between nutritional hazard and disease outcomes sometimes allow COI and DALY/HALY be applied to nutritional hazards. Risk assessment methodology can be applied to chemical hazards and microbiological hazards, when it is known as quantitative microbiological risk assessment (QMRA). Although the same procedure is followed, the calculations and the information required are quite different. Both RA types aim to calculate human exposure to a particular food safety hazard - the chemical contaminant and the pathogen, respectively – through food consumption. The main difference is that MRA calculates the pathogenic contamination of food at time of consumption and numbers of people getting ill from consuming that food, whereas chemical RA calculate the exposure of the contaminant by food at the time of consumption and evaluate if this exposure is below or above the Tolerable Daily Intake (ADI), or similar. For ranking several chemical contaminants in food at once, methods typically applied are the risk ratio method and the scoring method. These methods either multiply or divide a parameter for occurrence of the chemical (e.g. concentration) and the severity of the hazard (e.g. TDI).

MCDA was mostly applied to rank microbiological hazards, but could also be applied for ranking chemical hazards, or both. However, when applied to ranking two or even three types of hazards (if nutritional hazards are included), great care must be taken in designing the MCDA so that a common set of parameters are identified which are relevant to all hazard groups.

For some methods, such as risk matrix and risk ratio, essential data needs appear to be smaller than with other methods, like RA, CRA and MCDA. However, it is more that these former methods could also be applied when less information is available, although ideally larger amounts would be available. This is in contrast to the latter methods that have a large demand of quantitative data and can only be applied when these data are available. When new, additional data become available, this should be processed by the method selected in order to update risk ranking results. Automatic or easy updating of results is an issue that was hardly touched upon in the risk ranking method application found in literature, but this issue merits further investigation. In addition, automatic or easy updating of results could also be used for the scenario analyses or sensitivity analyses of results. It requires an IT application of data, stored in datasheets or databases, linked to model calculations expressed in scripts. Methods most suitable for such an automatic update are RA, risk ratio, risk scoring, risk matrices, COI, HALY, and MCDA. It is more difficult to apply with CRA, WTP and expert synthesis. For WTP and expert synthesis, the context in which participants make their choices will be altered (e.g. changes in relative prices or perceived risk), and hence primary data will need to be collected again with the method designed to reflect the altered context.

Methods that apply quantitative approaches demand more data and result in more precise outcomes with a better description of the uncertainties, assuming that data quality is high. Qualitative methods can be used when data are scarce, e.g., when emerging hazards, such as botanicals, are to be ranked. They also have the advantage of generating rich descriptive material, by which insights into the reasoning behind the opinions (or ranking decisions) of participants can be obtained. In the cases of limited data availability, the appropriate methods are risk matrix, flow charts/decision trees with an emphasis on input from experts, or a ranking based solely on expert synthesis of available quantitative and qualitative information. In the cases of the latter, use qualitative inputs, the outcomes will also be less precise.

In general, quantitative methods taking into account uncertainty and variability require more time and resource than qualitative methods. However, most methods that are used for qualitative situations can also be used semi-quantitatively or quantitatively. And in the latter case, they would also require an equal amount of time and resource. For instance, risk matrices and expert judgements can be used in a simple application using qualitative input or asking the expert to provide their qualitative opinion, respectively. When performed more quantitatively also expert judgement and risk matrices are also resource intensive.

In principle, all methods can account for uncertainty and variability in the input data used, acknowledging this information is more precise and quantitatively defined with the quantitative methods. RA and CRA, both of which can accommodate uncertainty and variability in the input data, appear to be very useful methods for providing quantitative results, provided their substantial data requirements are met. . Semi-quantitative and qualitative methods could also allow for inclusion of uncertainty. Two methods do not have the capacity to consider uncertainty in terms of outcomes, these being risk matrix and flow/decision charts.

Risk ranking can be based on a narrow range of parameters, e.g., measurements of exposure and effect on human health, such as risk ratio or the scoring method, or can include wider issues such as economic impacts and societal preferences. Most methods are demanding of time and other resources, e.g., for primary data collection, although some predefined tools for risk ranking are openly available . MCDA is typically applied when, besides exposure and effect, other metrics need to be considered, such as the consumers' perception of risk associated with different hazards. The strength of this method is in this wider applicability and the involvement of stakeholder groups to assess preference functions and weights. It is often applied in a multi-stakeholder situation. WTP is typically applied when consumer perception on food safety is to be included in the risk ranking.

The results of risk rankings should be interpreted carefully as relatively small differences in methodology can result in changes in final rankings. There is a need for transparency regarding the method used and its application and adequate explanation so users can understand the rationale which has been used to derive the numbers.

An important element of all risk ranking activities is communication of the outputs to interested end-users, including the general public. A question arises as to how such communication processes are developed from the outputs of these different risk ranking methodologies in forms which are both understandable and relevant to different interested end-user communities, and there is no comparative analysis currently available. Including risk perceptions may, for example, increase the relevance of the outputs to the general public, but the extent to which such communication is trusted compared to the communication of outputs from risk ranking methodologies where this has not been the case requires further research, as does the development of a more general communication strategy regarding risk ranking practices and allocation of resources to associated risk mitigation activities.

In conclusion, this study showed there is a wide range of methods that can be used for ranking food related hazards, based on their impact on human health. It has demonstrated that there is no single best risk ranking method. Each of the method categories has its own strengths and weaknesses. The most suitable methods should be selected based on the risk manager's requirements and needs, as well as available resources , the risk ranking task at hand, data availability and the characteristics of the methods. To this end, close communication between risk managers and risk assessors is needed to identify to the most suitable method for risk ranking. Uncertainties associated with data input need to be clearly stated. To date, this is not part of the standard procedure of most methods. This overview is valuable for industrial and governmental risk managers, and risk assessors for selecting the most appropriate methods for risk ranking of food and diet related hazards on the basis of human health impact. The overview will facilitate this decision process and allow for a structured and transparent selection of the most appropriate risk ranking method.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the grant received from EFSA for performing this study (EFSA/SCOM/01/2013) as well as critical reflections from Arianna Chiusolo (EFSA) and her colleagues on the study results.

References

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

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.

Batz MB, Hoffman S and Morris JG Jr, 2011. Ranking the Risks: The 10 Pathogen-Food Combinations With The Greatest Burden on Public Health. Emerging Pathogens Institute, University of Florida, Gainsville, USA, 7 pp..

Batz MB, Hoffmann S, Morris JG Jr., and Clenn, J, 2012. Ranking the disease burden of 14 pathogens in food sources in the United States using attribution data from outbreak investigations and expert elicitation. [Erratum appears in J Food Prot. 2012 Aug;75(8):1366]. 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.

Bouwknegt M, Friesema IHM, Van Pelt W and Havelaar AH, 2013. Disese burden of food-related pathogens in the Netherlands, 2011. Institute of Public Health and the Environment (RIVM), Bilthoven, the Netherlands. 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. US Department of Agriculture, US. Agricultural Economic Report No AER-741, 93 pp.

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.

Chen Y, Dennis SB, Hartnett E, Paoli G, Pouillot R, Ruthman T and Wilson M, 2013. FDA-iRISK - A Comparative Risk Assessment System for Evaluating and Ranking Food-Hazard Pairs: Case Studies on Microbial Hazards. Journal of Food Protection, 76, 376-385.

Codex Alimentarius, 2001. Codex Alimentarius Commission - Procedural Manual - Twelfth Edition. Joint FAO/WHO Food Standards Programme, FAO, Rome, Italy, 175 pp

Codex Alimentarius, 2014. Principles and guidelines for the conduct of microbiological risk assessment CAC/GL 30-1999. Adopted 1999, with amendments 2012, 2014. FAO/WHO, 5 pp. Available at: http/www.codexalimentarius.org/input/download/.../CXG_030e_**2014**.pdfCope 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.

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.

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, 2010. Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal, 8, 1637.

EFSA, 2011. Overview of the procedures currently used at EFSA for the assessment of dietary exposure to different chemical substances. EFSA Journal, 9.

EFSA, 2012a. Scientific Opinion on Reflecting on the experiences and lessons learnt from modelling on biological hazards. EFSA Journal, 10, 2725.

EFSA, 2012b. Scientific Opinion on the development of a risk ranking framework on biological hazards. EFSA Journal, 10, 2724.

EFSA, 2012c. Scientific Opinion on the public health hazards to be covered by inspection of meat (poultry). EFSA Journal, 20, 2741-2920.

EFSA, 2012d. Scientific Opinion on the public health hazards to be covered by inspection of meat (swine). EFSA Journal, 20, 2351-2549.

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.

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

FAO/WHO, 2008. Microbiological hazards in fresh leafy vegetables and herbs. Meeting Report Microbiological risk assessment series 14, Rome, Italy. Available at:ftp://ftp.fao.org/docrep/fao/011/i0452e/i0452e00.pdf, Accessed 28 November 2015.

FAO and WHO, 2012. Multicriteria-based ranking for risk management of foodborne parasites. Report of a joint FAO/WHO expert meeting, 3-7 september 2012, FAO Headquarters, Rome, Italy, 47 pp.

Fazil A, Rajic A, Sanchez J, and McEwen S, 2008. Choices, choices: The application of multi-criteria decision analysis to a food safety decision making problem. Journal of Food Protection, 71, 2323-2333.

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, Hobart, Australia. Available at: http://www.foodsafetycentre.com.au/riskranger.php, Accessed 19 September 2015.

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

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.

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.

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

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, 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.. Washington DC, USA: Resources for the future, 39 pp.

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

Howard GA, Ferorze M F, Teunis P, Mahmud SG, Davison A, and 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.

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.National Institute of Public Health and the Environment, Bilthoven, the Netherlands.

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.

Lim SSV, T.; Flaxman, A. D.; Danaei, G.; Shibuya, K.; Adair-Rohani, H.; Amann, M.; Anderson, H. R.; Andrews, K. G.; Aryee, M.; Atkinson, C.; Bacchus, L. J.; Bahalim, A. N.; Balakrishnan, K.; Balmes, J.; Barker-Collo, S.; Baxter, A.; Bell, M. L.; Blore, J. D.; Blyth, F.; Bonner, C.; Borges, G.; Bourne, R.; Boussinesq, M.; Brauer, M.; Brooks, P.; Bruce, N. G.; Brunekreef, B.; Bryan-Hancock, C.; Bucello, C.; Buchbinder, R.; Bull, F.; Burnett, R. T.; Byers, T. E.; Calabria, B.; Carapetis, J.; Carnahan, E.; Chafe, Z.; Charlson, F.; Chen, H.; Chen, J. S.; Cheng, A. T. A.; Child, J. C.; Cohen, A.; Colson, K. E.; Cowie, B. C.; Darby, S.; Darling, S.; Davis, A.; Degenhardt, L.; Dentener, F.; Des Jarlais, D. C.; Devries, K.; Dherani, M.; Ding, E. L.; Dorsey, E. R.; Driscoll, T.; Edmond, K.; Ali, S. E.; Engell, R. E.; Erwin, P. J.; Fahimi, S.; Falder, G.; Farzadfar, F.; Ferrari, A.; Finucane, M. M.; Flaxman, S.; Fowkes, F. G. R.; Freedman, G.; Freeman, M. K.; Gakidou, E.; Ghosh, S.; Giovannucci, E.; Gmel, G.; Graham, K.; Grainger, R.; Grant, B.; Gunnell, D.; Gutierrez, H. R.; Hall, W.; Hoek, H. W.; Hogan, A.; Hosgood Iii, H. D.; Hoy, D.; Hu, H.; Hubbell, B. J.; Hutchings, S. J.; Ibeanusi, S. E.; Jacklyn, G. L.; Jasrasaria, R.; Jonas, J. B.; Kan, H.; Kanis, J. A.; Kassebaum, N.; Kawakami, N.; Khang, Y. H.; Khatibzadeh, S.; Khoo, J. P.; Kok, C.; Laden, F.; Lalloo, R.; Lan, Q.; Lathlean, T.; Leasher, J. L.; Leigh, J.; Li, Y.; Lin, J. K.; Lipshultz, S. E.; 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, G. A.; Merriman, T. R.; Micha, R.; Michaud, C.; Mishra, V.; Hanafiah, K. M.; Mokdad, A. A.; Morawska, L.; Mozaffarian, D.; Murphy, T.;

Formatted: English (U.S.)

Naghavi, M.; Neal, B.; Nelson, P. K.; Nolla, J. M.; Norman, R.; Olives, C.; Omer, S. B.; Orchard, J.; Osborne, R.; Ostro, B.; Page, A.; Pandey, K. D.; Parry, C. D. H.; Passmore, E.; Patra, J.; Pearce, N.; Pelizzari, P. M.; Petzold, M.; Phillips, M. R.; Pope, D.; Pope Iii, C. A.; Powles, J.; Rao, M.; Razavi, H.; Rehfuess, E. A.; Rehm, J. T.; Ritz, B.; Rivara, F. P.; Roberts, T.; Robinson, C.; Rodriguez-Portales, J. A.; Romieu, I.; Room, R.; Rosenfeld, L. C.; Roy, A.; Rushton, L.; Salomon, J. A.; Sampson, U.; Sanchez-Riera, L.; Sanman, E.; Sapkota, A.; Seedat, S.; Shi, P.; Shield, K.; Shivakoti, R.; Singh, G. M.; Sleet, D. A.; Smith, E.; Smith, K. R.; Stapelberg, N. J. C.; Steenland, K.; Stöckl, H.; Stovner, L. J.; Straif, K.; Straney, L.; Thurston, G. D.; Tran, J. H.; Van Dingenen, R.; Van Donkelaar, A.; Veerman, J. L.; Vijayakumar, L.; Weintraub, R.; Weissman, M. M.; White, R. A.; Whiteford, H.; Wiersma, S. T.; Wilkinson, J. D.; Williams, H. C.; Williams, W.; Wilson, N.; Woolf, A. D.; Yip, P.; Zielinski, J. M.; Lopez, A. D.; Murray, C. J. L.; 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.

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.

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, Plass D and Kretzschmar MEE, 2014. Estimating the current and future burden of communicable diseases in the European Union and EEA/EFTA. Institute of Public Health and the Environment (RIVM), Bilthoven, the Netherlands. Report 210474001/2014.

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.

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

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.

Newsome RT, Tran N, Paoli GM , Jaykus LA, Tompkin B, Miliotis M, Ruthman T, Hartnett E, Busta FF , Petersen B, Shank F, McEntire J, Hotchkiss J, Wagner M, and Schaffner DW, 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.

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.

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.

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 and 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.

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.

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.

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.

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.

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 Kreijl CF, Knaap AGAC and Van Raaij JMA, 2006. Our food, our health - Healthy diet and safe foods in the Netherlands (in Dutch). Institute of Public Health and the Environment (RIVM), Bilthoven, the Netherlands. Report 270555009, 364 pp.

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

Formatted: Dutch (Netherlands)

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.

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.

LEGENDS TO FIGURES

Figure 1: Framework for risk ranking of chemicals, adapted from Bu et al. (2013).

Figure 2: Example of Risk matrix

Figure 1.

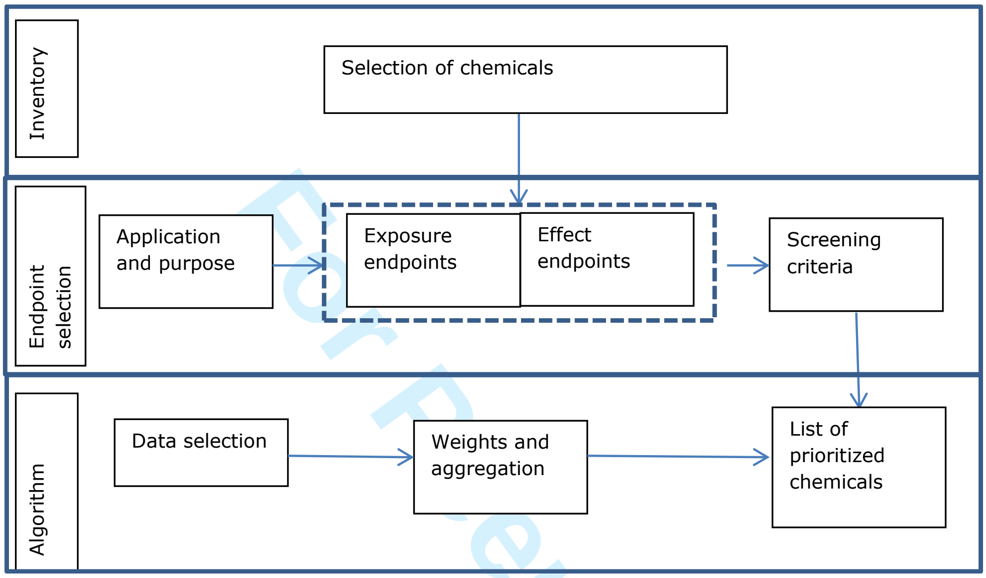

Inventory
Selection of chemicals
Endpoint selection
Application and purpose
Exposure endpoints
Effect endpoints
Screening criteria
Algorithm
Data selection
Weights and aggregation
List of prioritized chemicals

Figure 2

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

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

Type hazard/field	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
Health adjusted live years	90	13	9	98	18
Socio-economic methods	3296	47	15	3366	20

Table 2: Number of references per method categories for risk ranking of the food and/or nutritional hazards

Type hazard	Risk assessment	Comparative risk assessment	Ratio	Scoring	Cost of illness	HALY	Stated preference[1]	MCDA[1]	Risk Matrix	Flow chart / Decision trees	Expert synthesis
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	3	1	0	0	1[4]	0	1	0	2	0
Other	0	0	0	0	0	0	1	1	0	0	1
Sum	95	3	38	24	10	29	8	19	16	22	15

[1]WTP: Willingness to Pay; HALY; health adjusted live years, MCDA: Multi Criteria Decision Analyses;

[2]One reference described both chemical and microbiological hazards;

[3]Three references described both chemical and microbiological hazards;

[4]One reference described both chemical and nutritional hazards.

ANNEX 1. Literature search protocol

a) Search strategy and search strings

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 for methods for risk ranking and prioritisation in the field of food related issues (step 1), then narrowing down the methods relating to size of anticipated impact on human health (step 2), and finally focusing on chemical hazards, biological hazards, nutritional components, or social issues related to food (step 3). The strategy steps and final search strings are as follows:

Step 1: Captured titles/subject headings that studied methods and tools for risk ranking and prioritization related to food issues. This step included the following search strings:

TOPIC = (risk*[i] 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 = (food* OR agri* or agro*OR environ*) 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:

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)

TOPIC = ("socio* impact" OR "econ* impact" OR WTP OR cost* OR WTA)

Step 3: Captured titles/subject headings that investigated specific application fields of biological hazards, chemical hazards, nutritional components in food, or social science issues related to food hazards, from consumer and governance perspectives. This step included the following search strings:

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*

OR:

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*

OR

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*

OR

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*

b) Evaluation criteria

The references judged to be relevant for the study objectives were evaluated for eligibility and quality of the described research. References were included when:

1. Reference was relevant for the objective of the literature review;
   - References discussing prioritisation/ranking methods for human health risks and/or,
   - References describing risk prioritization/ranking methods applied for environmental/ecological risks and/or,
   - References to risk prioritization, risk analysis, risk assessment methods and/or risk modelling included in abstract and/or,
   - Any relevance of the work for application to human health, including references on drinking water and/or,
   - Abstract indicates socio-economic research methodology is employed.
2. Reference came from international peer-reviewed journals;
3. Methods in the reference were well described, (semi-)quantitative or qualitative, user-friendly, transparent, structured, and objective;
4. Methods in the reference 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.

Criteria for excluding references were:

- References discussing only parts of a method (only exposure or only human health effects), 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.

Table 3. Characteristics of risk ranking methods related to food safety

Characteristic	Risk Assessment	Comparative Risk Assessment	Ratio (Exposure/ Effect)	Scoring method	Cost of Illness	HALY[1]	WTP[1]	MCDA[1]	Risk Matrix	Flow charts /Decision trees	Expert Synthesis
Amount of resources (time, money)	High	High	Moderate	Moderate	Moderate	Moderate	High	High	Low	Low	Moderate /Low
Level of output	Quantitative	Quantitative	Semi-quantitative	Semi-quantitative	(Semi-) quantitative	(Semi-) quantitative	(Semi-) quantitative	Semi-quantitative	Qualitative/semi-quantitative	Qualitative	Qualitative
Easy to explain to stakeholders (laymen)?	No	No	Yes	Yes	Yes	No	No	Yes	Yes	No	Yes
Inclusion stakeholder perception	Not possible	Not possible	Not possible	Possible	Not possible	Not possible	Possible	Possible	Not possible	Possible	Possible
Inclusion uncertainty	Possible	Possible	Possible	Possible	Possible	Possible	Possible	Possible	Not possible	Not possible	Possible
Inclusion weights for the risk ranking criteria	Not possible	Not possible	Not possible	Possible	Not possible	Not possible	Not possible	Possible	Not possible	Not possible	Possible
Inclusion human incidences	Possible	Possible	Not possible	Not possible	Possible	Possible	Possible	Possible	Not possible	Possible	Possible
Inclusion economic impact	Not possible	Not possible	Not possible	Not possible	Possible	Not possible	Possible	Possible	Not possible	Possible	Possible
Common method of communication (in addition to reports)	Graphs/Tables	Graphs/Tables	Tables	Tables	Graphs/Tables	Graphs/Tables	Graphs/Tables	Graphs/Tables	Graphs	Decision Tree	Tables
Essential data needed											
Human incidence data needed?	No	Yes	No	No	Yes	Yes	Yes	No	No	No	No
Dose-response data needed?	Yes	Yes	No	No	No	No	No	No	No	No	No
Occurrence data (concentration, prevalence, dose) needed?	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No
Food consumption data needed?	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No
Growth models needed (only applicable for microbiological hazards)?	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No

Toxicological reference values (ADI, TDI etc) needed (only applicable for chemical hazards)?	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No

[1]WTP: Willingness to Pay; HALY; health adjusted live years, MCDA: Multi Criteria Decision Analysis

RESEARCH PAPER

Critical review of methods for risk ranking of food related hazards, based on risks for human health

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], M. D'Agostino[4], D. Coles[2], <u>H. J. P. Marvin[1]</u> L. J. Frewer[2], ~~H. J.P. Marvin[1]~~

[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; [3]Technical University of Denmark, National Food Institute, Morkhoj Bygade 19, 2860 Soborg, Denmark; [4]Food and Environmental Research Agency, Sand Hutton, York, North Yorkshire, YO41, 1 LZ.

ABSTRACT

This study aimed to critically review methods for ranking risks related to food safety and dietary hazards on the basis of their anticipated human health impacts. A systematic literature review was performed to identify and characterize methods for risk ranking from the fields of food, environmental science and socio-economic sciences. The review used a predefined search protocol, and covered the bibliographic databases Scopus, CAB Abstracts, Web of Sciences, and PubMed over the period 1993-2013.

All references deemed relevant, on the basis of of predefined evaluation criteria, were included in the review, and the risk ranking method characterized. The methods were then clustered – based on their characteristics - into eleven method categories. These categories included: risk assessment, comparative risk assessment, risk ratio method, scoring method, cost of illness, health adjusted life years, multi-criteria decision analysis, risk matrix, flow charts/decision trees, stated preference techniques and expert synthesis. Method categories were described by their characteristics, weaknesses and strengths, data resources, and fields of applications.

It was concluded there is no single best method for risk ranking. The method to be used should be selected on the basis of risk manager/assessor requirements, data availability, and the characteristics of the method. Recommendations for future use and application are provided.

KEY-WORDS

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

1. INTRODUCTION

Ranking of health risks related to food safety and nutrition is generally recognised as the basis for risk-based priority setting and resource allocation. It permits governmental and regulatory organisations to allocate their resources efficiently to the most significant public health problems (Van Kreijl et al., 2006). Within the area of food, risk is defined as the analysis and prioritization of the combined probability of food contamination, consumer exposure and the size of the anticipated public health impact of specific chemical, microbiological and/or nutritional hazards related to food. It is the combination of the *probability that a hazard may occur* in a food product and the *effect of exposure to the hazard on human health* (Codex Alimentarius 2001). Risk ranking has been applied to food safety monitoring programs and has shown to increase the efficiency of monitoring and to decrease inspection costs, both in practice and from theoretical calculations (Baptista et al., 2012; Presi et al., 2008; Reist et al., 2012).

To date, various risk ranking methods are available that prioritise food safety risks (Van Asselt et al., 2012). Methods vary from qualitative, through semi-quantitative, to quantitative methods (Cope et al., 2010; Van Asselt et al., 2012). ~~Examples of tools that apply quantitative methods are the swift QMRA tool (Evers and Chardon, 2010) and iRISK, which is a comparative risk assessment system for evaluating and ranking food hazard pairs (Chen et al. 2013, see http://www.foodrisk.org). As quantitative methods can be very elaborate, semi-quantitative tools such as Risk Ranger (Ross and Sumner, 2002) have also been developed (Food Safety Centre, 2010).~~ Most methods are based on the 'technical' concept of risk being a function of presence of the hazard and severity of its impact on human health. However, some methods also involve other metrics, which may be considered in decision making, e.g., consumer perceptions of risk. In order to determine which methods are most suitable for ranking food related risks, it is important to follow a structured, objective and transparent approach to identifying and evaluating the available methods (van Asselt et al., 2013).

The aim of the current study was to review available methods for ranking risks associated with food on the basis of anticipated health impact, to characterize the methods and to provide recommendations for their use.

2. MATERIAL AND METHODS

2.1 Protocol for literature review

A literature review was conducted which aimed to identify risk ranking methodologies that can be used to prioritize food related hazards, on the basis of the size of anticipated health impact. Hazards are defined as those agents that can be present in food and can negatively affect human health (Codex Alimentarius, 2001). Hazards included in this study were nutritional, chemical and microbiological hazards. The review covered methods from the fields of natural/life (food) science, socio-economic sciences and food safety governance, published during the period 1993-2013. Risk ranking methods from fields outside food science (i.e. environmental sciences and socio-economic methods) were also included to evaluate their appropriateness for application in food science. The literature review followed the principles of a systematic literature review as described by EFSA (2010). A protocol for the structured literature review was defined *a priori*, including search strings and criteria for evaluation of the literature references (Annex 1).

2.2 Literature review

<u>Review methodology</u>

a. Scientific articles were identified using the following bibliographic databases: Web of Science, Scopus, PubMed, and CAB Abstracts. In addition, the general search engine Google was used to search for reports, (the ‘grey literature’), from relevant international and national organisations, authorities, and agencies (e.g., EFSA, EMA, WHO/FAO, FDA, Health Canada, OECD). The literature search focused on papers and reports published in English.
b. The set of search strings was applied leading to an initial set of search results. All retrieved references were stored in an Endnote database. Duplicates, a result of using four different bibliographic databases, were removed.

c. The references resulting from the initial set of search results were screened for their relevance to the study objectives by applying the 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 allocated to one of three categories and placed in the corresponding category of the Endnote database:

- *Relevant for this study*: the reference was included;
- *Possibly relevant for this study*: uncertain if the reference was relevant for the study;
- *Not relevant for this study*: the reference was determined to be out of scope.

An inter-observer check was conducted with a randomly selected subset (10%) of both selected and excluded references.

d. In tier 2, the full text of the references that were in the *Relevant* and *Possibly relevant* groups of the Endnote database were retrieved. By reading the full texts, the papers/reports were evaluated for their relevance to the field of interest and their quality using the evaluation criteria. When deemed relevant, the reference was retained or moved to the group *Relevant* in the Endnote database. When deemed not relevant, the reference was moved to the group *Not relevant* in the Endnote database. Also at this stage, an inter-observer check was conducted; certain (randomly chosen) literature references were evaluated by two experts from ~~of~~ the team (from different disciplines) in order to gain insights into the variation between the evaluation results of two different experts.

e. Citations used in the reports/references of the final Endnote database were screened for additional relevant references, published after 1993 (snowball citation), and steps c) and d) were applied to them.

Evaluation of references

For each reference stored in the *Relevant* category of the Endnote database, the risk ranking method and its characteristics were evaluated in depth. A summary of the information obtained was stored in an excel sheet, using a unique row for each reference. The format of the excel sheet was defined beforehand, starting from the template developed by EFSA's BIOHAZ panel (EFSA, 2012b), but with some modification to increase relevance to the objectives of the current study. Separate columns were utilised for information about the reference (author names, title, abstract, journal, volume and page numbers), and for storing the results from the critical evaluation of the risk ranking methods including: the type of tool (short description); field of application (microbiological, chemical, and/or nutritional hazards); 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 requirements that describe the model variables (e.g., human population data, or microbial numbers); method of data collection, describing how the necessary data were collected and which data sources were used, and finally data integration, describing how data were integrated in the application described in the reference. Based on this evaluation, the references and the evaluated methods were categorised into different groups of methods. The method categories were then described according to the following characteristics: scope, application area, approach, strengths and weaknesses, and perspective for use ~~by~~ by risk managers~~stakeholders~~. At this stage, reviews on risk ranking methods and other relevant literature were also consulted..

3. RESULTS ~~AND DISCUSSION~~

3.1 Literature search

At tier 1, application of the search strings and removal of duplicates led to the retrieval of the following numbers of references (Table 1): 6021 for chemical/toxicological hazards; 2932 for microbiological hazards; 1049 for nutritional hazards; 112 references using health adjusted live years method; and 3358 references using socio-economic methodology. The latter two method groups were considered since they could potentially include each of the three types of hazards (microbiological, chemical and/or nutritional hazards). The total numbers of references appearing in tier 2 are somewhat higher than in tier 1 due to snowballing citations. In total 253 references were judged to be relevant.

3.2 Description of risk ranking methods

Based on the evaluation of the methods described in the relevant references, the risk ranking methods were classified, according to methodology, into the following categories: 1) Risk Assessment (RA), 2) Comparative risk assessment (CRA), 3) Risk ratio method, 4) Scoring method, 5) Risk matrix, 6) Flow charts (including decision trees and influence diagrams), 7) Cost of illness (CoI), 8) Health adjusted life years (HALY), 9) Multi criteria decision analysis (MCDA), 10) Stated preference methods, and 11) Expert judgement. Table 2 shows the numbers of references that presented a particular method category, per type of hazard. All methods included both presence of the hazard and its severity~~exposure and effect~~. Method categories differed in ~~, although~~ the way in which these two factors were evaluated and combined to come to an estimate of the risk~~covered varied between the method categories~~. In some instances, a combination of methods was applied, in which case the study was classified to its main category.

RA was by far the most frequently applied method. This method was applied to both chemical and microbiological hazards. For each of the chemical and microbiological hazards, about one third of all tier 1 references described the application of a RA to a particular hazard. However, as the procedure for each of the chemical and microbiological RA is comparable, only references describing guidelines for performing a RA were included. Risk ratio, scoring, risk matrices and flow charts were mostly applied to chemical hazards, whereas CoI, HALY, and expert judgments were mostly used for ranking microbiological hazards (Table 2). Ranking methods for nutritional hazards were fewer, and were mostly based on RA, CRA and expert judgement (Table 2). CRA, CoI, and stated preferences were the methods that were applied least frequently, with CRA used in three studies about~~-~~ nutritional hazards, and the latter two methods primarily applied to microbiological hazards. A few studies have considered both chemical and microbiological hazards in their ranking, applying methods for CoI and HALY. Summaries of each method and characteristics are presented in the following sections and in Table 3.

3.2.1. Risk Assessment

Scope: A RA for a chemical or microbiological hazard aims to estimate the risk for human health associated with the presence of the hazard in one or more food products, and total food consumption. Numerous risk assessments have been applied to chemical and microbiological hazards in food. WHO (WHO, 2009) and Codex Alimentarius (2014~~2012~~) have provided guidelines regarding the principles and methods for the risk assessment of chemical contaminants and pathogens in foods. Although the application of the RA methodology is tailored to the hazard type, the principles for performing a risk assessment for both types of hazards are identical, consisting of the following four steps: hazard identification, exposure assessment, hazard characterisation, and risk characterization.

Application area: Risk assessment is usually applied for one identified (chemical or microbiological) hazard occurring in a specific food commodity and for a predefined population, with the purpose of characterizing~~-~~ the associated health risk. Apart from this, an important reason for conducting a RA is to evaluate the impact of control measures to reduce the risk. If the results of different RA are compared (e.g. for different hazards or different foods), the RA can be used for risk ranking.

Approach: Various RA approaches for chemical and microbiological hazards in food were identified, applying different combinations of deterministic, probabilistic (or stochastic), qualitative, semi-quantitative, and quantitative modelling. Furthermore, different approaches were used for the exposure assessment and the hazard characterization steps. EFSA (2011) published an overview of procedures for current RA methods for dietary exposure of different chemical substances. The need for development of harmonized approaches, and future exploration of cumulative exposure assessments, is identified. In 2012, EFSA published its experiences gained with Quantitative Microbiological Risk Assessment (QMRA) studies (EFSA, 2012a).

Strengths and weaknesses: In RA, all available scientific and technical information and data, as well as variability and uncertainties are systematically organized and analysed. It is a well-structured method, providing insights into what is known and what is not known. In particular, RA offers the opportunity to address uncertainties in a transparent way, e.g., *via* sensitivity analyses and/or modelling and simulation runs. It could be the most precise method to estimate risks, including the relevant uncertainties. However, a RA for one chemical or microbiological hazard usually requires a lot of time, data and knowledge. Ranking risks related to various hazards in food using outcomes of individual RAs will take even more resources and RAs are often hampered by a lack of quantitative data. Lack of data, selection of models to fit to the data, and assumptions that need to be made give rise to uncertainties in the outcomes. Recently, several tools for relative risk assessment for pathogens of pathogen-food combinations have been published. Examples of such tools applying quantitative methods are the swift QMRA tool (Evers and Chardon, 2010) and iRISK, which is a relative risk assessment system for evaluating and ranking food-hazard pairs (Chen et al. 2013, see http:// https://irisk.foodrisk.org/). An example of a semi-quantitative approach is Risk Ranger (Ross and Sumner, 2002) developed by Food Safety Centre (2010).

Perspective for use by risk manager: Applied optimally, RA should disseminate key information regarding risk from exposure to food hazards to policy makers, decision makers and the public. RA are very useful for providing insights into gaps in knowledge and issues associated with high levels of uncertainty. However, they may not be suitable for risk ranking given the large amounts of data, knowledge and resources needed.

3.2.2. Comparative risk assessment

Scope: A Comparative Risk Assessment (CRA) analysis can estimate the number of deaths that would be prevented in a given period if current distributions of risk factor exposure were changed to a hypothetical alternative distribution (Danaei et al., 2009; Micha et al., 2012). In these papers, CRA is restricted to comparisons of deaths and it is, therefore, not comparable to a risk assessment or a relative risk assessment.

Application area: Three applications of CRA have been found; each of them studied the impact of dietary factors on disease mortality. 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. Lim and co-workers (2012) investigated burden of disease and injury attributable to 67 risk factors (including chemical hazards and nutritional imbalances) in 21 regions through application of a systematic analysis for the Global

Burden of Disease Study 2010. Although a CRA analysis as described below was not performed by Lim et al. (2012), several elements of a CRA analysis were included.

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. Input needed to determine the PAF include: a) effect size (relative risk estimate) of the causal diet-disease relationship, b) optimal or theoretical minimum-risk exposure distribution, c) dietary risk factor exposure distribution in the population and, d) total number of disease-specific deaths (plus non-fatal events, when available) in the population. Data sources for obtaining these inputs include epidemiological studies, systematic reviews, meta-analysis, nationally representative nutrition surveys and mortality databases.

Strengths and weaknesses: A CRA analysis is a systematic assessment of unbiased data collected in national and international surveys as well as the peer reviewed literature. It allows for consistent, comparable and quantitative assessment of the global impact of risk factors on disease by sex- and age-specific groups. A CRA analysis requires knowledge and resources (manpower, money, data), which makes it expensive to perform. Unbiased data are also needed, e.g., to establish exposure distributions or causal diet-disease relationships, which may often not be easily accessible or available. The weights of different diseases are not considered. Uncertainties associated with a CRA analysis can be high because of data limitations.

Perspectives for use by risk manager~~stakeholders~~: A CRA analysis offers a global assessment of the impact of dietary factors on disease mortality, which is very valuable for priority setting and policy making. However, with large and overlapping uncertainty ranges for the different risk factors, ranking of modifiable dietary risk factors may be difficult.

3.2.3. Risk ratio method

Scope: Risk ratios or quotients refer to a~~re~~ quantitative method in which ~~derived by dividing~~ estimates of exposure are divided by estimates of effect. For this purpose, data are needed regarding the amounts of the hazard consumed (either the dose or the concentration) as well as a measure for the effect of the hazards that are studied.

Application: The risk ratio method has usually been applied to rapidly screen the risk of a range of chemical compounds in order to rank them. Most studies applied the method to rank pesticides, although five studies focused on microbiological hazards, and one study applied the method to rank both chemical and microbiological hazards.

Approach: For chemical contaminants, 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) (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 (Bench Mark 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 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). Applications of the method to microbiological hazards used different criteria, such as costs and effective dose.

Strengths and weaknesses: This method is easy to understand, and can be applied once concentration data and toxicological reference values are available; it only needs an estimate for both amounts of the hazardous material consumed and the effect of the hazard on human health. For emerging chemical hazards, e.g., nanomaterials, toxicological reference values are usually not available. ~~Furthermore, concentration data are also not always available. It may thus be difficult to rank all hazards of interest due to data limitations~~.

Perspectives for use by risk manager ~~stakeholders~~: The method can give a quick answer on the risk of food safety hazards for human health, and can be applied to both chemical and microbiological hazards.

3.2.4. Scoring method~~s~~

Scope: This method is based on semi-quantitative scoring of both exposure and effect of the hazard on human health, followed by their multiplication (or – in one reference - addition).

Application: Scoring methods provide a simple risk ranking method to 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).

Approach: When a scoring method is applied, both exposure and severity (or effect) endpoints are considered. However, endpoints for exposure and effect can vary. Various endpoints have been used to estimate exposure, such as chemical transformation properties (degradability, half-life), mobility/distribution (such as bioaccumulation factors (BAF) or bioconcentration factors (BCF)), release, 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, selection of appropriate endpoints for a specific study is one of the steps in ranking risks according to a scoring method. Examples of endpoints for effect on human health might include acute toxicity, carcinogenicity, or reproductive toxicity, and can be based on LD50, MOAEL, BMDL10 etc. Once criteria are set, endpoints are classified semi-quantitatively, e.g., using scores from 1 to 3 or from 1 to 5, as applied in, for example ~~e.g., (~~Penrose et al. ~~, (~~(1994). After this classification system for endpoints has been established, data sources need to be found in order to assign scores for exposure and effect. These sources can be based on literature, available data and/or expert opinion. –Scores subsequently need to be aggregated, which is mainly done by multiplying exposure and effect (see, e.g., 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 employ 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., 2005~~l~~). 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.

Strengths and weaknesses: This semi-quantitative method is easy to conduct once scores have been assigned to the model variables. Furthermore, it allows the inclusion of stakeholder perceptions in assigning the scorings a~~s~~nd the importance (to each stakeholder) of each model variable is reflected by the weighting allocated to it. The assigned weights should then be clearly documented to guarantee a transparent approach.

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

3.2.5. Risk matrices

Scope: Just like the scoring methods, risk matrices also make use of scoring both exposure and effect endpoints. The difference between scoring methods and risk matrices is that, in the latter, the exposure and effect endpoints are not aggregated by multiplication or addition, but are 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 or microbiological hazards for which limited quantitative data are available. This method has, for example, been applied for ranking the risks of nanomaterials (O'Brien and Cummins, 2011; Sorensen et al., 2010; Zalk et al., 2009).

Approach: Both the likelihood of occurrence and the consequences of the hazard for human health are scored into one of several classes; see Figure 2 for an example. Classes that could be used for likelihood of occurrence are: almost certain, likely, possible, unlikely and rare. Classes that could be used for the consequences are: insignificant, minor, moderate, major and severe. The division into these classes is subjective. Then, risk classes are assigned to the combinations of Likelihood and Consequences, e.g., being L (low), M (moderate), H (high), and E (extreme), as shown in Figure 2. Risk classification may also be based on scores. Zalk et al. (2009), for example, classified nanomaterials based on scores for probability and severity, and the results were depicted in a risk matrix. The results can also be visualized using spider web plots, as conducted by, (e.g)., Ranke and Jastorff (2000), who classified various endpoints using scores from 1-4, and compared plots for the various compounds to obtain an indication of the most risky ones.

Strengths and weaknesses: The risk matrix method is qualitative or semi-quantitative, and thus less accurate than methods based on concentration data and dose-response relationships or toxicological reference values. It provides a visualisation for both presence of the hazard and its effects, giving direct insights into the way these two elements contribute to the overall risk of a hazard. For example, a hazard may present a high risk due to a high exposure, although its severity is low. Alternatively, due to its high toxicity, it may present a high risk rank despite low exposure. Matrices will give more information to the risk manager compared to other methods that produce a list of hazards according to the overall risk alone. However, the division between different categories for presence of the hazard (e.g. low, medium high occurrence) and its effects (e.g. low, medium, high toxicity) is subjective and, thus, other results are obtained when with other divisions.

Perspectives for use by risk manager: 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.

3.2.6. Flow charts

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

Application: Flow charts or decision trees can be used for various purposes. In general these methods are used to obtain a qualitative indication about the risks associated with hazards. Haase et al. (2012), for example, established a decision tree for nanoparticles to determine whether a full risk assessment is required or not. EFSA described guidelines for classifying chemical hazards as negligible, low, medium, and high risks (EFSA, 2012c, 2012d).

Approach: A flow chart is generally based on several questions that need to be answered in order to arrive at a certain risk class. Questions can be based on the likelihood that specific chemicals or microbiological hazards are present in the study object; evidence of occurrence or incorrect practice in the food chain, the toxicological profile, and the outcome of national monitoring programmes (EFSA, 2012c, 2012d). Eisenberg and McKone (1998) used a Classification and Regression Tree Algorithm (CART) to specify the chemical and environmental properties and Monte Carlo simulations to estimate human exposure. Schmidt et al. (2011) utilized a decision support system (DSS) to rank genetically modified organisms (GMOs), based on a decision tree and rules, indicators and baselines, and thresholds (such as the LD50) (Schmidt et al., 2011). DSS may also be combined with multi-criteria decision analysis (MCDA). Critto (2007), for example, utilised a DSS system to evaluate ecological observations and ecotoxicological tests for contaminated sites and then incorporated MCDA and expert judgments into the ranking. This approach might also be used for ranking food safety risks.

Strengths and weaknesses: Flow charts/decision trees present a straightforward method with clear questions for which only qualitative information is needed, although quantitative information can be used where available. The method can, thus, be used for a quick screening of food safety hazards, in order that the most relevant ones may subsequently be investigated in more detail. However, this method strongly depends on expert input and it is, therefore, essential to perform a rigorous expert elicitation study. Furthermore, this type of method is vulnerable to being less transparent than other methods, as it is not always clear why hazards end up being classified as a high, medium or low risk. Therefore, for each hazard 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 risk manager: It is important to set up the right questions for inclusion in a flow chart/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.

3.2.~~5~~7. Cost of Illness method

Scope: The underlying research objective of the Cost of Illness (CoI) approach is distinct from those of the methodologies described so far. CoI studies acquire data for conducting economic analysis in order to obtain a ranking in terms of how society might allocates scarce resources when addressing food-related hazards. The procedure involves ~~methodology implies~~ calculating the directs costs to society related to disease and death ~~in society~~ due to chemical, microbial and/or nutritional 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 loss of income. Subject to data availability, it is possible to compare large numbers of food risks.

Application area: 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).

Approach: 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 experiences each type of impact, defined as the disease severity class. A critical point is whether it is restricted to acute effects, or whether long-term effects (sequelae and deaths) 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).

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 (e.g., Batz et al., 2011) or frequency distributions can be used in Monte Carlo or stochastic simulation models (Lake et al., 2010; Kemmeren et al., 2006). The costs incurred at each state are calculated, often including the categories of direct health costs, indirect health costs, and indirect non-health costs.

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 one year, whereas other end-points, such as life-long disabilities, are strongly affected by discounting. Hence, the effect of discounting will differ per hazard (Kemmeren et al., 2006) and the rate of interest selected.

<u>Strengths and weaknesses</u>: The CoI method employs readily available and reliable data (Buzby et al., 1996) and the calculations are transparent and relatively simple. The same disease incidence data are used in HALY calculations so it is relatively efficient to produce both sets of rankings at the same time and they are, to some extent, complementary. A combined risk ranking can also be produced. 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, the measures excludes non-workers, and do not -address perceived quality of life including factors such as pain and stress (Golan et al., 2005). 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. The results are dependent on the assumptions made *inter alia* about medical outcomes and the prevailing labour market.

<u>Perspectives for use by risk manager~~stake-holders~~</u>: CoI is a well-tried technique with well-understood limitations relating to missing data, and failure of the approach 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 is ~~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.

<u>3.2.~~6~~8. Health adjusted life years (Burden of Disease)</u>

<u>Scope:</u> 'Health adjusted life years (HALY)' 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 is then multiplied by the duration of that health state. A descriptive summary of the various HALYs is presented by Mangen et al. (2014).

<u>Application area</u>: HALY measures 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, i.e., the burden of disease. HALY estimates such as disability adjusted life years (DALYs) or quality-adjusted life years

(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. 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 HALY approach has been applied to rank different pathogens and chemical contaminants in the same food category, 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: Data are required for estimating the number of cases with the most relevant types of acute illnesses, chronic sequelae and mortality (also termed health outcomes) arising from exposure to the hazards under consideration. Different types of hazards (chemical, microbiological or nutritional) require different types of data and modelling approaches (Crettaz et al., 2002; Hofstetter, 2002; Mangen et al., 2010; Mangen et al., 2014; Pennington et al., 2002), but after the final DALY/QALY calculations have been made, the risks estimates should be readily comparable. DALY/QALY estimates may also be included in several of the other risk ranking methods such as RA (Howard et al. (2007); Newsome et al. (2009)~~; Chen et al. (2013)~~), CRA (Lim et al. (2012)), MCDA (Ruzante et al. (2010)), risk matrixes, flow charts/decision trees or in expert syntheses.

Strengths and weaknesses: HALY methodologies readily allow 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 et al., 2012) and ECDCs initiative for developing methodologies for measuring current and future burden of communicable diseases (Mangen et al., 2014).

DALYs and QALYs are semi-quantitative estimates based on disability scoring, and their accuracy is highly dependent on the quality of input data and risk assessment models used for estimating the incidences of relevant health outcomes. In the applied studies, the methods for estimating the incidences of relevant health outcomes varied widely. The estimated DALY or QALY values seem to be relatively precise quantitative estimates, and there is a risk of over-interpretation of the relative differences, if the level of uncertainty is not addressed. A general methodological weakness is

Formatted: Dutch (Netherlands)
Formatted: Dutch (Netherlands)
Field Code Changed
Formatted: Dutch (Netherlands)
Formatted: Dutch (Netherlands)
Field Code Changed
Formatted: Dutch (Netherlands)
Field Code Changed
Formatted: Dutch (Netherlands)
Formatted: Dutch (Netherlands)
Formatted: Dutch (Netherlands)
Field Code Changed
Formatted: Dutch (Netherlands)
Formatted: Dutch (Netherlands)
Field Code Changed
Formatted: Dutch (Netherlands)
Formatted: Dutch (Netherlands)
Field Code Changed
Formatted: Dutch (Netherlands)
Field Code Changed
Formatted: Dutch (Netherlands)
Formatted: Dutch (Netherlands)
Formatted: Dutch (Netherlands)
Field Code Changed

inadequate evidence to estimate the incidences of chronic disability, especially in cases with few or no symptoms during the acute phase of a disease. Another methodological weakness is that the concept of DALYs assumes a continuum from good health to disease, disability, and death which is independent of time – a concept not universally accepted. Also, stakeholders have difficulty to understand the concept and what is meant by it.

Perspectives for use by risk manager~~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 RA or models for estimation of reported cases are available, the resources needed to estimate DALYs are moderate. However, development of RA 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 of the ‘societal production losses’ related to the included hazards. If HALY estimates from different studies are to be used in risk ranking, then differences in the methodology employed and the comparability of the studies must be considered. For monitoring purposes, risk ranking models estimating HALYs 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).

~~3.2.7. 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 for risk reduction. 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.~~

~~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 safety risks is limited and largely confined to valuing individual disease reduction measures or comparing alternative risk management options within single food-disease problem, see e.g., Mørkbak & Nordström (2009) and Miller et al. (2005). Golan et al (2005) concluded that, at~~

present,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: Using stated preference methods, 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. In contrast to the CoI approach, stated preference methods include the value individuals place on other factors for which no markets exist such as, for instance, (not) experiencing pain. 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. One of the stated preference methods, willingness to pay (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). As an example, Mørkbak and Nordström (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. This approach defines the choices which individuals make in terms of the levels of key attributes (such as high/low price, probability of illness etc) which are associated with each of the goods being compared.

Strengths and weaknesses: 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 revealing which societal groups (e.g., by age, background) ranks a particular risk most highly (see Haninger and Hammitt (2011) for an example). The aggregated value of benefits (or societal WTP) of food safety (e.g., reduced risks) can be compared with the costs for achieving them since both costs and benefits are expressed in monetary units.

However, WTP is a difficult technique to apply, and is prone to errors and bias unless conducted meticulously. Experience so far has been in comparing only 2 to 4 alternative risks. 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. Moreover, 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.

Perspectives for use by stakeholders. These techniques provide a means to incorporate societal preferences in ranking and decision making. However, experience in the food safety field as yet is only modest, and there is scope to develop techniques still further.

3.2.89. Multi-Criteria Decision Analysis (MCDA)

Scope: MCDA is an approach which has the potential to evaluate multiple - often conflicting - criteria in decision making. It allows for comparison of different risks on common basis, by simultaneous consideration of provides a fairly transparent means of identifying the salient parameters of a problem (technical information, uncertainty and different stakeholder preferences), , and can potentially include both quantitative and qualitative data, and and the integration of large amounts of complex information. to allow for comparison of different risks on a common basis. MCDA has a long history of use in various decision contexts, e.g., in nanomaterial risk assessment. 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 enable making leading to more informed and better decisions. In the context of risk ranking, important criteria utilized in food safety can be identified through a process of expert or lay consultation, which may include not only public health impacts but also perception, costs – an in case of interventions – also weight of evidence, and practicality associated with the interventions

Application area: MCDA can be applied to any range of problems, which can be defined in terms of a common set of criteria. As the scientifically 'best' solution may be inadequate in terms of acceptability to society, utilize resources which or not available, or be sub-optimal in terms of allocating resources, stakeholder methods are sometimes used to capture the preferences of consumers, citizens and/or experts. Hence, stakeholder engagement can feature in MCDA in particular when politically acceptable solutions are to be defined. Indeed, MCDA which combines expert judgement across a range of relevant criteria appears to be the second most popular method for relative risk ranking of microbiological hazards, after RA.

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 (Fazil et al., 2008), can be supplied with technical information in relation to each risk criterion to assist their deliberations. The selection of preference functions and weights are an integral and core part of the MCDA methodology and must be selected when conducting a risk ranking. An example is provided by Ruzante et al. (2010) who utilized the method to develop a prioritization framework for foodborne risks that considered not only public health impacts but also market impact, consumer risk acceptance and perception, and social sensitivity. For each risk under consideration, participants give each criterion either a numerical score or an ordinal ranking such as 'high', medium' and 'low'. In an MCDA, a key issue that could differentiate the possible approaches is whether weights are applied to criterion scores and, if so, how they are elicited. At the simplest level, , criteria could be considered as equal, which, however, may resuly in the oversimplification of experts' views. Alternatively, experts can allocate weights for each MCDA criterion, thereby indicating the degree of importance they put on each criterion in the MCDA outputs. The weighted scores are then combined to produce a single score for eachissue, permitting scores to be ranked. Another well-known example of a MCDA method for ranking pathogen-produce combinations is the Pathogen-Produce Pair Attribution Risk Ranking Tool (P^{3}ARRT) developed by FDA (Anderson et al., 2011), which is available free (http://foodrisk.org/exclusives/rrt). Fazil et al. (2008) applied MCDA for the ranking of food safety interventions, considering amongst others cost, effectiveness, and weight of evidence. MCDA methods and applications vary in their complexity; they may even allow for

Formatted: Superscript

probabilistic modelling and sensitivity analyses. Recently, alternative methods for performing a MCDA have been developed and employed, e.g., by Havelaar et al. (2010), in order to minimise the biases linked with experts' direct weighting of the MCDA criteria.

Strengths and weaknesses: MCDA allows consideration of stakeholder perceptions by using the weights and preference functions they assign to the various criteria in the analysis. Furthermore, economic impact or other criteria that are deemed relevant can be included, in addition to human health criteria. 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, therefore, allows inclusion of subjective elements that may also be important for risk managers to include in their decision making processes, depending on the aim of the ranking exercise. Alternative scenarios using weights and preference functions for various input factors can be compared. However, MCDA outcomes are more difficult to communicate compared t~~op~~ more straightforward methods such as risk matrices or scoring methods, as various criteria are included, which are weighted and prioritized ~~often each having~~ differently~~ weights~~. Furthermore, this method needs expert or stakeholder input in order to derive the weights and preference functions for the criteria. Therefore this method has weaknesses that are linked to the elicitation of information from experts (see below), i.e., the 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); the need to consider how to elicit experts' own uncertainties regarding their views, opinions, judgments; and - last but not least – the need to consider possible ways to combine individual opinions without masking variability in the experts' views.

Perspectives for use by risk manager~~stakeholders~~: This systematic method is very valuable in cases where stakeholder perceptions are required to be included in the risk ranking, as ~~a~~ weight~~sing~~s and preference functions can be assigned to the various model variables. This method also allows the inclusion of factors other than effect and exposure endpoints, e.g. from the social-economic field, or in terms of policy development, which makes it a very versatile tool. The application of MCDA will provide a single number for ranking. However, the underlying calculations can be difficult for the non-expert to understand for those without expertise in the methodology~~grasp~~.

3.2.10. 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 for risk reduction. 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.

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 safety risks is limited and largely confined to valuing individual disease reduction measures or comparing alternative risk management options within single food-disease problem, see e.g., Mørkbak & Nordström (2009) and Miller et al. (2005). Golan et al (2005) concluded that, at present,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: Using stated preference methods, 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. In contrast to the CoI approach, stated preference methods include the value individuals place on other factors for which no markets exist such as, for instance, (not) experiencing pain. 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.

One of the stated preference methods, willingness to pay (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). As an example, Mørkbak and Nordström (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. This approach defines the choices which individuals make in terms of the levels of key attributes (such as high/low price, probability of illness etc) which are associated with each of the goods being compared.

Strengths and weaknesses: 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 revealing which societal groups (e.g., by age, background) ranks a particular risk most highly (see Haninger and Hammitt (2011) for an example). The aggregated value of benefits (or societal WTP) of food safety (e.g., reduced risks) can be compared with the costs for achieving them since both costs and benefits are expressed in monetary units.

However, WTP is a difficult technique to apply, and is prone to errors and bias unless conducted meticulously. Experience so far has been in comparing only 2 to 4 alternative risks. 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. Moreover, 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.

Perspectives for use by risk manager. These techniques provide a means to incorporate societal preferences in ranking and decision making. However, experience in the food safety field as yet is only modest, and there is scope to develop techniques still further.

3.2.9. Risk matrices

Scope: Just like the scoring methods, risk matrices also make use of scoring both exposure and effect endpoints. The difference between scoring methods and risk matrices is that, in the latter, the exposure and effect endpoints are not aggregated by multiplication or addition, but are 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 or microbiological hazards for which limited quantitative data are available. This method has, for example, been applied for ranking the risks of nanomaterials (O'Brien and Cummins, 2011; Sorensen et al., 2010; Zalk et al., 2009).

Approach: Both the likelihood of occurrence and the consequences of the hazard for human health are scored into one of several classes; see Figure 2 for an example. Classes that could be used for likelihood of occurrence are: almost certain, likely, possible, unlikely and rare. Classes that could be used for the consequences are: insignificant, minor, moderate, major and severe. Then, risk classes are assigned to the combinations of Likelihood and Consequences, e.g., being L (low), M (moderate), H (high), and E (extreme), as shown in Figure 2. Risk classification may also be based on scores. Zalk et al. (2009), for example, classified nanomaterials based on scores for probability and severity, and the results were depicted in a risk matrix. The results can also be visualized using spider web plots, as conducted by, (e.g)., Ranke and Jastorff (2000), who classified various endpoints using scores from 1-4, and compared plots for the various compounds to obtain an indication of the most risky ones.

Strengths and weaknesses: The risk matrix method is qualitative or semi-quantitative, and thus less accurate than methods based on concentration data and dose-response relationships or toxicological reference values. It provides a visualisation for both effect and exposure of the hazard, giving direct insights into the way these two elements contribute to the overall risk of a hazard. For example, a hazard may present a high risk due to a high exposure, although its severity is low. lAternatively, due to its high toxicity, it may present a high risk rank despite low exposure. Matrices will give more information to the risk manager compared to other methods that produce a list of hazards according to the overall risk alone. However, the classification for consequences and likelihood may not be fully underpinned by the available data. Furthermore, the method depends on expert input, requiring a rigorous 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.

3.2.10. Flow charts

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

Application: Flow charts or decision trees can be used for various purposes. In general these methods are used to obtain a qualitative indication about the risks associated with hazards. Haase et al. (2012), for example, established a decision tree for nanoparticles to determine whether a full risk assessment is required or not. EFSA described guidelines for classifying chemical hazards as negligible, low, medium, and high risks (EFSA, 2012c, 2012d).

Approach: A flow chart is generally based on several questions that need to be answered in order to arrive at a certain risk class. Questions can be based on the likelihood that specific chemicals or microbiological hazards are present in the study object; evidence of occurrence or incorrect practice in the food chain, the toxicological profile, and the outcome of national monitoring programmes (EFSA, 2012c, 2012d). Eisenberg and McKone (1998) used a Classification and Regression Tree Algorithm (CART) to specify the chemical and environmental properties and Monte Carlo simulations to estimate human exposure. Schmidt et al. (2011) utilized a decision support system (DSS) to rank genetically modified organisms (GMOs), based on a decision tree and rules, indicators and baselines, and thresholds (such as the LD50) (Schmidt et al., 2011). DSS may also be combined with multi-criteria decision analysis (MCDA). Critto (2007), for example, utilised a DSS system to evaluate ecological observations and ecotoxicological tests for contaminated sites and then incorporated MCDA and expert judgments into the ranking. This approach might also be used for ranking food safety risks.

Strengths and weaknesses: Flow charts/decision trees present a straightforward method with clear questions for which only qualitative information is needed, although quantitative information can be used where available. The method can, thus, be used for a quick screening of food safety hazards, in order that the most relevant ones may subsequently be investigated in more detail. However, this method strongly depends on expert input and it is, therefore, essential to perform a rigorous expert

~~elicitation study. Furthermore, this type of method is vulnerable to being less transparent than other methods, as it is not always clear why hazards end up being classified as a high, medium or low risk. Therefore, for each hazard 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 flow chart/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.~~

3.2.11. Expert judgement

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.

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). The inclusion of public perceptions, priorities and values may result in a different ranking being reached to that derived from using scientific experts alone. This 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 (DeKay et al., 2005).

Approaches: A variety of methods is available, for application in workshops or in surveys, which may be characterised by the flows of information which take place between the participants and the research team (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, e.g., MCDA and the Carnegie-Mellon approach. In MCDA ~~these~~

~~approaches~~, the judgement of stakeholders is used ~~to rank risks and~~ to allocate weights and potentially also on the way to weight the different criteria and in establishing the preferences to the different attributes whereas the Carnegie-Mellon approach produces risk rankings. ~~to produce a multi-attribute ranking~~. Approaches also 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. Four approaches were identified:

- 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). Commonly used approaches are workshops and the Classical Delphi method (Van der Fels-Klerx et al., 2002).
- Survey based on existing knowledge of lay or expert participants (i.e. minimal technical communication during the study), as applied by, e.g., Schwarzinger et al. (2010) and Harrington (1994).
- 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. FAO/WHO, 2008), 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). 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).
- ~~-~~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. 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 (DeKay et al., 2005).

Strengths and weaknesses: Judgement-based methods provide additional information to that of technical assessments, e.g., when 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. However, judgement-based methods require a 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 ned for meticulous preparation the method is 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 risk manager ~~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. Also, they could provide a means of incorporating societal values into risk ranking.

DISCUSSION AND CONCLUSIONS

A ~~systematic~~ literature review has been performed on methodologies for ranking risks related to chemical, microbiological and nutritional hazards in food, on the basis of their anticipated effects on human health. The results showed that a range of risk ranking methodologies has been applied depending on the purpose of the specific study. They ~~various methods~~ have been grouped into eleven main categories, determined primarily by the type(s) of hazard that can be ranked, data needs, and uncertainty. Some methods allow ranking of different hazards types (chemical, microbiological), whereas others allow ranking only within one hazard category.

Four of the eleven method groups can be applied to all three types of hazards (microbiological, chemical and nutritional), either alone or in combination, these being MCDA, risk matrices, stated preferences techniques, and expert synthesis. For microbiological hazards, there is a close relationship between exposure and resulting levels of illness and death, which allows CoI and DALY/HALY calculations to be made. With chemical contamination of food, there is no such direct relationship between the contamination and resulting diseases/deaths in the population, since effects on human health are long-term and, hence, the cause-effect relationship is difficult to establish. Consequently, these methods is are not often applied to chemical food contamination, although an exception is the study by Kemmeren et al. (2006) who calculated DALYs for chemical contaminants, using assumptions on the relations between chemical food contamination and disease outcomes. Although health effects of nutritional hazards are often evident only in the longer term, recent improved availability of insights from long-term epidemiological studies on the cause-relationships between nutritional hazard and disease outcomes sometimes allow –COI and DALY/HALY be applied to nutritional hazards. Risk assessment methodology can be applied to chemical hazards and microbiological hazards, when it is known as quantitative microbiological risk assessment (QMRA). Although the same procedure is followed, the calculations and the information required are quite different. Both RA types aim to calculate human exposure to a particular food safety hazard - the chemical contaminant and the pathogen, respectively – through food consumption. The main difference is that MRA calculates the pathogenic contamination of food at time of consumption and numbers of people getting ill from consuming that food, whereas chemical RA calculate the exposure of the contaminant by food at the time of consumption and evaluate if this exposure is below or above the Tolerable Daily Intake (ADI), or similar. For ranking several chemical contaminants in food at once, methods typically applied are the risk ratio method and the scoring method. These methods either multiply or divide a parameter for occurrence of the chemical (e.g. concentration) and the severity of the hazard (e.g. TDI).

MCDA was mostly applied to rank microbiological hazards, but could also be applied for ranking chemical hazards, or both. However, when applied to ranking two or even three types of hazards (if nutritional hazards are included), great care must be taken in designing the MCDA so that a common set of parameters are identified which are relevant to all hazard groups.

For some methods, such as risk matrix and risk ratio, essential data needs appear to be smaller than with other methods, like RA, CRA and MCDA. However, it is more that these former methods could also be applied when less information is available, although ideally larger amounts would be available. This is in contrast to the latter methods that have a large demand of quantitative data and can only be applied when these data are available. When new, additional data become available, this should be processed by the method selected in order to update risk ranking results. Automatic or easy updating of results is an issue that was hardly touched upon in the risk ranking method application found in literature, but this issue merits further investigation. In addition, automatic or easy updating of results could also be used for the scenario analyses or sensitivity analyses of results. It requires an IT application of data, stored in datasheets or databases, linked to model calculations expressed in scripts. Methods most suitable for such an automatic update are RA, risk ratio, risk scoring, risk matrices, COI, HALY, and MCDA. It is more difficult to apply with CRA, WTP and expert synthesis. For WTP and expert synthesis, the context in which participants make their choices will be altered (e.g. changes in relative prices or perceived risk), and hence primary data will need to be collected again with the method designed to reflect the altered context.

Methods that apply quantitative approaches demand more data and result in more precise outcomes with a better description of the uncertainties, assuming that data quality is high. Qualitative methods can be used when data are scarce, e.g., when emerging hazards, such as botanicals, are to be ranked. They also have the advantage of generating rich descriptive material, by which insights into the reasoning behind the opinions (or ranking decisions) of participants can be obtained. In the cases of limited data availability, the appropriate methods are risk matrix, flow charts/decision trees with an emphasis on input from experts, or a ranking based solely on expert synthesis of available quantitative and qualitative information. In the cases of the latter, use qualitative inputs, the outcomes will also be less precise.

In general, quantitative methods taking into account uncertainty and variability require more time and resource than qualitative methods. However, most methods that are used for qualitative situations can also be used semi-quantitatively ly or quantitatively. And in the latter case, they would also require an equal amount of time and resource. For instance, risk matrices and expert judgements can be used in a simple application using qualitative input or asking the expert to provide their qualitative opinion, respectively. When performed more quantitatively also expert judgement and risk matrices are also resource intensive.

In principle, all methods can account for uncertainty and variability in the input data used, .Aacknowledging this information is more precise and quantitatively defined with the quantitative methods. RA and CRA, both of which can accommodate uncertainty and variability in the input data, appear to be very useful methods for providing quantitative results, provided their substantial data requirements are met. In general, methods that apply quantitative approaches demand more resources and result into more precise outcomes with a better description of the uncertainties. Semi-quantitative and qualitative methods could also allow for inclusion of uncertainty. Two methods do not have the capacity to consider uncertainty in terms of outcomes, these being risk matrix and flow/decision charts. Some methods allow ranking of different hazards types (chemical, microbiological), whereas others allow ranking only within one hazard category.

RA and CRA, both of which can accomodate uncertainty and variability in the input data, appear to be very useful methods for providing quantitative results, provided their substantial data requirements are met. More qualitative methods could be used when data are scarce, e.g., when emerging hazards, such as botanicals, are to be ranked. In the cases of limited data availability, the appropriate methods are MCDA, risk matrix, flow charts/decision trees with an emphasis on input from experts, or a ranking based solely on expert judgement.

Risk ranking can be based on a narrow range of parameters, e.g., measurements of exposure and effect on human health, such as risk ratio or the scoring method, or can include wider issues such as economic impacts and societal preferences. Most methods are demanding of time and other resources, e.g., for primary data collection, although some predefined tools for risk ranking are openly available exist. MCDA is typically applied when, besides exposure and effect, other metrics need to be considered, such as the consumers' perception of risk associated with different hazards. The strength of this method is in this wider applicability and the involvement of stakeholder groups to assess preference functions and weights. It is often applied in a multi-stakeholder situation. WTP is typically applied when consumer perception on food safety is to be included in the risk ranking.

The results of risk rankings should be interpreted carefully as relatively small differences in methodology can result in changes in final rankings. There is a need for transparency regarding the method used and its application and adequate explanation so users can understand the rationale which has been used to derive the numbers.

An important element of all risk ranking activities is communication of the outputs to interested end-users, including the general public. A question arises as to how such communication processes are developed from the outputs of these different risk ranking methodologies in forms which are both understandable and relevant to different interested end-user communities, and there is no comparative analysis currently available. Including risk perceptions may, for example, increase the relevance of the outputs to the general public, but the extent to which such communication is trusted compared to the communication of outputs from risk ranking methodologies where this has not been the case requires further research, as does the development of a more general communication strategy regarding risk ranking practices and allocation of resources to associated risk mitigation activities.

.

In conclusion, Tthis study showed there is a wide range of methods that can be used for ranking food related hazards, based on their impact on human health. It has demonstrated that there is no single best risk ranking method. Each of the method categories has its own strengths and weaknesses. The most suitable methods should be selected based on the risk manager's requirements and needs, (as well as available resources) , the risk ranking task at hand, data availability and the characteristics of the methods. To this end, close communication between risk managers and risk assessors is needed to identify- to the most suitable method for risk ranking. Uncertainties associated with data input need to be clearly stated. To date, this is not part of the standard procedure of most methods.

This overview is valuable for industrial and governmental risk managers, and risk assessors for selecting the most appropriate methods for risk ranking of food and diet related hazards on the basis of human health impact. The overview will facilitate this decision process and allow for a structured and transparent selection of the most appropriate risk ranking method.

ACKNOWLEDGDMENTS

The authors gratefully acknowledge the grant received from EFSA for performing this study (EFSA/SCOM/01/2013) as well as critical reflections from Arianna Chiusolo (EFSA) and her colleagues on the study results.

References

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

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.

Batz MB, Hoffman S and Morris JG ~~Jr~~ Jr~~G~~, 2011. Ranking the Risks: The 10 Pathogen-Food Combi nations With The Greatest Burden on Public Health. ~~In (pp. 70 p.):~~ Emerging Pathogens Institute, University of Florida, Gainsville, USA, 7 pp..

Batz MB, Hoffmann S, ~~and~~ Morris JG~~,~~ Jr., and Clenn, J, 2012. Ranking the disease burden of 14 pathogens in food sources in the United States using attribution data from outbreak investigations and expert elicitation. [Erratum appears in J Food Prot. 2012 Aug;75(8):1366]. 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.

Bouwknegt M, Friesema IHM, Van Pelt W and Havelaar AH, 2013. Disese burden of food-related pathogens in the Netherlands, 2011. Institute of Public Health and the Environment (RIVM), Bilthoven, the Netherlands. ~~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, US. Agricultural Economic Report No AER-741, 93 pp.

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.

Chen Y, Dennis SB, Hartnett E, Paoli G, Pouillot R, Ruthman T and Wilson M, 2013. FDA-iRISK - A Comparative Risk Assessment System for Evaluating and Ranking Food-Hazard Pairs: Case Studies on Microbial Hazards. Journal of Food Protection, 76, 376-385.

Codex Alimentarius, 2001. Codex Alimentarius Commission - Procedural Manual - Twelfth Edition. Joint FAO/WHO Food Standards Programme, FAO, Rome, Italy, 175 pp

Codex Alimentarius, 2014 ~~2012~~. Principles and guidelines for the conduct of microbiological risk assessment CAC/GL 30-1999. Adopted 1999, with amendments 2012, 2014. FAO/WHO, 5 pp. Available at: http/
www.codexalimentarius.org/input/download/.../CXG_030e_**2014**.pdf

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.

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.

Formatted: Superscript

Formatted: Indent: Left: 0 cm, First line: 0 cm

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.

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, 2010. Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal, 8, 1637.

EFSA, 2011. Overview of the procedures currently used at EFSA for the assessment of dietary exposure to different chemical substances. EFSA Journal, 9.

EFSA, 2012a. Scientific Opinion on Reflecting on the experiences and lessons learnt from modelling on biological hazards. EFSA Journal, 10, 2725.

EFSA, 2012b. Scientific Opinion on the development of a risk ranking framework on biological hazards. EFSA Journal, 10, 2724.

EFSA, 2012c. Scientific Opinion on the public health hazards to be covered by inspection of meat (poultry). EFSA Journal, 20, 2741-2920.

EFSA, 2012d. Scientific Opinion on the public health hazards to be covered by inspection of meat (swine). EFSA Journal, 20, 2351-2549.

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.

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

FAO/WHO, 2008. Microbiological hazards in fresh leafy vegetables and herbs. Meeting Report Microbiological risk assessment series 14, Rome, Italy. Available at:ftp://ftp.fao.org/docrep/fao/011/i0452e/i0452e00.pdf, Accessed 28 November 2015.

FAO and WHO, 2012. Multicriteria-based ranking for risk management of foodborne parasites. Report of a joint FAO/WHO expert meeting, 3-7 september 2012, FAO Headquarters, ~~In.~~ Rome, Italy, 47 pp~~: FAO headquarters~~.

Fazil A, Rajic A, Sanchez J, and McEwen S, 2008. Choices, choices: The application of multi-criteria decision analysis to a food safety decision making problem. Journal of Food Protection, 71, 2323-2333.

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, Hobart, Australia. Available at: http://www.foodsafetycentre.com.au/riskranger.php, Accessed 19 September 2015.

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

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.

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.

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

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, 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, 39 pp.

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

Howard GA, Ferorze M~~.~~ F, ~~.;~~ Teunis~~,~~ P, ~~.;~~ Mahmud~~,~~ S~~.~~ G, ~~.;~~ Davison A~~Annette, D.;~~ , and 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.

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:~~ National Institute of Public Health and the Environment, Bilthoven, the Netherlands.

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.

Lim SSV, T.; Flaxman, A. D.; Danaei, G.; Shibuya, K.; Adair-Rohani, H.; Amann, M.; Anderson, H. R.; Andrews, K. G.; Aryee, M.; Atkinson, C.; Bacchus, L. J.; Bahalim, A. N.; Balakrishnan, K.; Balmes, J.; Barker-Collo, S.; Baxter, A.; Bell, M. L.; Blore, J. D.; Blyth, F.; Bonner, C.; Borges, G.; Bourne, R.; Boussinesq, M.; Brauer, M.; Brooks, P.; Bruce, N. G.; Brunekreef, B.; Bryan-Hancock, C.; Bucello, C.; Buchbinder, R.; Bull, F.; Burnett, R. T.; Byers, T. E.; Calabria, B.; Carapetis, J.; Carnahan, E.; Chafe, Z.; Charlson, F.; Chen, H.; Chen, J. S.; Cheng, A. T. A.; Child, J. C.; Cohen, A.; Colson, K. E.; Cowie, B. C.; Darby, S.; Darling, S.; Davis, A.; Degenhardt, L.; Dentener, F.; Des Jarlais, D. C.; Devries, K.; Dherani, M.; Ding, E. L.; Dorsey, E. R.; Driscoll, T.; Edmond, K.; Ali, S. E.; Engell, R. E.; Erwin, P. J.; Fahimi, S.; Falder, G.; Farzadfar, F.; Ferrari, A.; Finucane, M. M.; Flaxman, S.; Fowkes, F. G. R.; Freedman, G.; Freeman, M. K.; Gakidou, E.; Ghosh, S.; Giovannucci, E.; Gmel, G.; Graham, K.; Grainger, R.; Grant, B.; Gunnell, D.; Gutierrez, H. R.; Hall, W.; Hoek, H. W.; Hogan, A.; Hosgood Iii, H. D.; Hoy, D.; Hu, H.; Hubbell, B. J.; Hutchings, S. J.; Ibeanusi, S. E.; Jacklyn, G. L.; Jasrasaria, R.; Jonas, J. B.; Kan, H.; Kanis, J. A.; Kassebaum, N.; Kawakami, N.; Khang, Y. H.; Khatibzadeh, S.; Khoo, J. P.; Kok, C.; Laden, F.; Lalloo, R.; Lan, Q.; Lathlean, T.; Leasher, J. L.; Leigh, J.; Li, Y.; Lin, J. K.; Lipshultz, S. E.; 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, G. A.; Merriman, T. R.; Micha, R.; Michaud, C.; Mishra, V.; Hanafiah, K. M.; Mokdad, A. A.; Morawska, L.; Mozaffarian, D.; Murphy, T.;

Formatted: English (U.K.)

Formatted: English (U.K.)

Formatted: English (U.K.)

Naghavi, M.; Neal, B.; Nelson, P. K.; Nolla, J. M.; Norman, R.; Olives, C.; Omer, S. B.; Orchard, J.; Osborne, R.; Ostro, B.; Page, A.; Pandey, K. D.; Parry, C. D. H.; Passmore, E.; Patra, J.; Pearce, N.; Pelizzari, P. M.; Petzold, M.; Phillips, M. R.; Pope, D.; Pope Iii, C. A.; Powles, J.; Rao, M.; Razavi, H.; Rehfuess, E. A.; Rehm, J. T.; Ritz, B.; Rivara, F. P.; Roberts, T.; Robinson, C.; Rodriguez-Portales, J. A.; Romieu, I.; Room, R.; Rosenfeld, L. C.; Roy, A.; Rushton, L.; Salomon, J. A.; Sampson, U.; Sanchez-Riera, L.; Sanman, E.; Sapkota, A.; Seedat, S.; Shi, P.; Shield, K.; Shivakoti, R.; Singh, G. M.; Sleet, D. A.; Smith, E.; Smith, K. R.; Stapelberg, N. J. C.; Steenland, K.; Stöckl, H.; Stovner, L. J.; Straif, K.; Straney, L.; Thurston, G. D.; Tran, J. H.; Van Dingenen, R.; Van Donkelaar, A.; Veerman, J. L.; Vijayakumar, L.; Weintraub, R.; Weissman, M. M.; White, R. A.; Whiteford, H.; Wiersma, S. T.; Wilkinson, J. D.; Williams, H. C.; Williams, W.; Wilson, N.; Woolf, A. D.; Yip, P.; Zielinski, J. M.; Lopez, A. D.; Murray, C. J. L.; 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.

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.

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. Institute of Public Health and the Environment (RIVM), Bilthoven, the Netherlands. ~~In RIVM~~ Report 210474001/2014.

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.

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

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.

Newsome RT, Tran N~~.;~~, Paoli GM ~~, G. M.;~~, Jaykus LA, ~~L. A.;~~ Tompkin~~,~~ B~~.;~~, Miliotis~~,~~ M~~.;~~, Ruthman~~,~~ T~~.;~~, Hartnett~~,~~ E~~.;~~, Busta~~,~~ FF~~. F.;~~, Petersen ~~,~~ B~~.;~~, Shank~~,~~ F~~.;~~, McEntire~~,~~ J~~.;~~, Hotchkiss~~,~~ J~~.;~~, Wagner~~,~~ M~~.;~~, and Schaffner DW~~, 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.

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.

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.

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

Rowe <u>G</u> and Frewer <u>LJ,</u> 2005. <u>A typology of public engagement mechanisms. Science, Technology and Human Values 30, 251-290.</u>

Formatted: Indent: Left: 0 cm, First line: 0 cm

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.

~~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.

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.

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.

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.

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 Kreijl CF, Knaap AGAC and Van Raiij JMA, 2006. Our food, our health - Healthy diet and safe foods in the Netherlands (in Dutch). Institute of Public Health and the Environment (RIVM), Bilthoven, the Netherlands. Report 270555009, ~~.~~364 pp.

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

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.~~

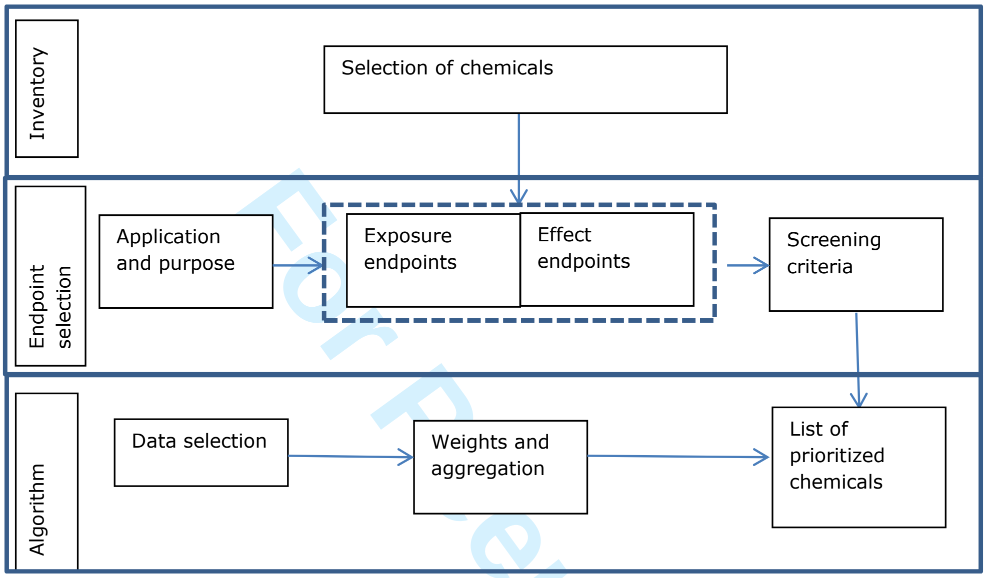
Zalk DM, Paik SY and Swuste P, 2009. Evaluating the Control Banding Nanotool: A qualitative risk assessment method for controlling nanoparticle exposures. Journal of Nanoparticle Research, 11, 1685-1704.

LEGENDS TO FIGURES

Figure 1: Framework for risk ranking of chemicals, adapted from Bu et al. (2013).

Figure 2: Example of Risk matrix

1228 Figure 1.



1229

1230

Figure 2

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

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

Type hazard/field	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
Health adjusted live years	90	13	9	98	18
Socio-economic methods	3296	47	15	3366	20

Table 2: Number of references per method categories for risk ranking of the food and/or nutritional hazards

Type hazard	Risk assessment	Comparative risk assessment	Ratio	Scoring	Cost of illness	HALY	Stated preference[1]	MCDA[1]	Risk Matrix	Flow chart / Decision trees	Expert synthesis
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	3	1	0	0	1[4]	0	1	0	2	0
Other	0	0	0	0	0	0	1	1	0	0	1
Sum	95	3	38	24	10	29	8	19	16	22	15

[1]WTP: Willingness to Pay; HALY; health adjusted live years, MCDA: Multi Criteria Decision Analyses;

[2]One reference described both chemical and microbiological hazards;

[3]Three references described both chemical and microbiological hazards;

[4]One reference described both chemical and nutritional hazards.

ANNEX 1. Literature search protocol

a) Search strategy and search strings

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 for methods for risk ranking and prioritisation in the field of food related issues (step 1), then narrowing down the methods relating to size of anticipated impact on human health (step 2), and finally focusing on chemical hazards, biological hazards, nutritional components, or social issues related to food (step 3). The strategy steps and final search strings are as follows:

Step 1: Captured titles/subject headings that studied methods and tools for risk ranking and prioritization related to food issues. This step included the following search strings:

TOPIC = (risk*[i] 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 = (food* OR agri* or agro*OR environ*) 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:

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)

TOPIC = ("socio* impact" OR "econ* impact" OR WTP OR cost* OR WTA)

Step 3: Captured titles/subject headings that investigated specific application fields of biological hazards, chemical hazards, nutritional components in food, or social science issues related to food hazards, from consumer and governance perspectives. This step included the following search strings:

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*

OR:

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*

OR

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*

OR

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*

b) Evaluation criteria

The references judged to be relevant for the study objectives were evaluated for eligibility and quality of the described research. References were included when:

1. Reference was relevant for the objective of the literature review;
    - References discussing prioritisation/ranking methods for human health risks and/or,
    - References describing risk prioritization/ranking methods applied for environmental/ecological risks and/or,
    - References to risk prioritization, risk analysis, risk assessment methods and/or risk modelling included in abstract and/or,
    - Any relevance of the work for application to human health, including references on drinking water and/or,
    - Abstract indicates socio-economic research methodology is employed.
2. Reference came from international peer-reviewed journals;
3. Methods in the reference were well described, (semi-)quantitative or qualitative, user-friendly, transparent, structured, and objective;
4. Methods in the reference 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.

Criteria for excluding references were:

- References discussing only parts of a method (only exposure or only human health effects), 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.

Table 3. Characteristics of risk ranking methods related to food safety

Characteristic	Risk Assessment	Comparative Risk Assessment	Ratio (Exposure/ Effect)	Scoring method	Cost of Illness	HALY[1]	WTP[1]	MCDA[1]	Risk Matrix	Flow charts /Decision trees	Expert Synthesis
Amount of resources (time, money)	High	High	Moderate	Moderate	Moderate	Moderate	High	~~Moderate~~High	Low	Low	Moderate /Low
Level of output	Quantitative	Quantitative	Semi-quantitative	Semi-quantitative	(Semi-) quantitative	(Semi-) quantitative	(Semi-) quantitative	Semi-quantitative	Qualitative/semi-quantitative	Qualitative	Qualitative
Easy to explain to stakeholders (laymen)?	No	No	Yes	Yes	Yes	No	No	Yes	Yes	No	Yes
Inclusion stakeholder perception	Not possible	Not possible	Not possible	Possible	Not possible	Not possible	Possible	Possible	Not possible	Possible	Possible
Inclusion uncertainty	Possible	Possible	Possible	Possible	Possible	Possible	Possible	Possible	Not possible	Not possible	Possible
Inclusion weights for the risk ranking criteria	Not possible	Not possible	Not possible	Possible	Not possible	Not possible	Not possible	Possible	Not possible	Not possible	Possible
Inclusion human incidences	Possible	Possible	Not possible	Not possible	Possible	Possible	Possible	Possible	Not possible	Possible	Possible
Inclusion economic impact	Not possible	Not possible	Not possible	Not possible	Possible	Not possible	Possible	Possible	Not possible	Possible	Possible
Common method of communication (in addition to reports)	Graphs/Tables	Graphs/Tables	Tables	Tables	Graphs/Tables	Graphs/Tables	Graphs/Tables	Graphs/Tables	Graphs	Decision Tree	Tables
Essential data needed~~DATA Needs~~											
Human incidence data needed?	No	Yes	No	No	Yes	Yes	Yes	No	No	No	No
Dose-response data needed?	Yes	Yes	No	No	No	No	No	No	No	No	No
Occurrence data (concentration, prevalence, dose) needed?	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No
Food consumption data needed?	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No
Growth models needed (only applicable for microbiological hazards)?	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No

Toxicological reference values (ADI, TDI etc) needed (only applicable for chemical hazards)?	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No

[1]WTP: Willingness to Pay; HALY; health adjusted live years, MCDA: Multi Criteria Decision Analysis