Part 2 Detailed Considerations

9. Qualitative: semi-quantitative: quantitative

9.1 Qualitative risk assessment

There are several examples of published qualitative risk assessments (e.g. Lake *et al.*, 2009; King, Lake and Cressey, 2011), although they tend to elicit less scientific attention than quantitative risk assessments.

It should be emphasized that the attributes of good risk assessment, as described in Chapter 3, apply equally to qualitative risk assessment. Appropriate data must be collected, documented and fully referenced and synthesized in a logical and transparent manner whichever method is employed.

Despite a number of large and well-publicized quantitative microbiological food safety risk assessment projects that have been completed, it is probable that the majority of risk assessments utilized by risk managers and policymakers in the fields of food safety, health and microbiology are not fully quantitative in the sense described in Chapter 3.

There may be a variety of reasons for this. Quantitative microbiological risk assessment is a specialized field and methods are still being developed, and the expertise and resources to complete them are not widely available. Equally, as noted earlier, the results of such assessments are not always 'accessible' to risk managers and other stakeholders. Thus, where a formal risk assessment (i.e. a body of work presented in a way that conforms to a set of risk assessment guidelines and specifically designed to estimate the magnitude of a risk) is commissioned by a risk manager, a qualitative risk assessment may be specified for reasons including:

- a perception that a qualitative risk assessment is much quicker and much simpler to complete;
- a perception that a qualitative risk assessment will be more accessible and easier for the risk manager or policymaker to understand and to explain to third parties;
- an actual or perceived lack of data, to the extent that the risk manager believes that a quantitative assessment will be impossible; or
- a lack of mathematical or computational skills and facilities for risk assessment, coupled with a lack of resources or desire to involve an alternative or additional source of expertise.

Whatever the reasons, many of them involve perceptions about the process of defensible qualitative risk assessment that, for reasons also mentioned above, are frequently not valid. Data and knowledge are required for any type of risk assessment, irrespective of whether qualitative, semi-quantitative or quantitative approaches are used. Numerical data are preferred, and a lack of appropriate crucial data will affect all approaches adversely. As data collection and documentation is usually the most time-consuming part of any risk assessment, and defensible logic is required to synthesize the data into an estimate or conclusion concerning the risk, a qualitative risk assessment will not necessarily be quicker or simpler to complete. In many cases, however, qualitative and semi-quantitative risk assessments are quicker to complete, and, whilst they require an equal degree of logic and considerable numeracy, they require fewer specialized mathematical and computational resources. A qualitative risk assessment has descriptions of the probability of an unwanted outcome in terms that are by their very nature subjective. It means that it is not necessarily easier either for the risk manager to understand the conclusions obtained from the risk assessment, or to explain them to a third party. Crucial to any formal risk assessment method is transparency, whether to describe how a numerical or a qualitative description of risk was achieved, because this enables

users to understand the basis of the assessment, to understand its strengths and limitations, to question or critique the assessment, or provide additional data or knowledge to improve the assessment. Additionally, because all approaches also require specialized medical, microbiological, biological, veterinary, epidemiological and other expertise, the inclusion of information and concepts from such a wide variety of areas of knowledge can make the risk assessment less accessible. Section 16.5 contains information about ways in which the results of risk assessment can be communicated.

9.1.1 *The value and uses of qualitative risk assessment*

Risk assessment, at its simplest, is any method that evaluates, or attempts to evaluate, a risk. Qualitative risk assessment is not, however, simply a literature review or description of all of the available information about a risk issue: it must also arrive at some conclusion about the probabilities of outcomes for a baseline risk and/or any reduction strategies that have been proposed. Both CAC (1999) and OIE (2018) state that qualitative and quantitative risk assessments have equal validity, though they did not specifically consider semi-quantitative risk assessment. However, neither organization explains the conditions under which qualitative and quantitative risk assessments are equally valid, and there is debate among risk experts about methods and approaches to be applied for qualitative risk assessment, and criteria for their validity. The World Trade Organization Committee on Sanitary and Phytosanitary Measures notes some advantages of quantitative expressions of risk (WTO, 2000):

> *"... quantitative terms, where feasible, to describe the appropriate level of protection can facilitate the identification of arbitrary or unjustified distinctions in levels deemed appropriate in different situations ... use of quantitative terms and/or common units can facilitate comparisons."*

However, when developing risk assessments, numerical results should be explained and put in context with a discussion of the limitations of the data and analysis, the important assumptions made, and the qualitative aspects of the risk not illuminated by quantitative analysis. The same underlying logic applies whether the assessment is quantitative or qualitative.

It is sometimes the case that a qualitative risk assessment is undertaken initially, with the intention of following up with a quantitative risk assessment if it is subsequently thought to be necessary or useful.

It may be the case that a qualitative assessment provides the risk manager or policymaker with all the information they require. For example, perhaps the information gathered includes some piece of evidence that shows that the risk is effectively indistinguishable from zero, and no more need currently be done. Or, conversely, perhaps evidence shows that it is obviously unacceptably large, or that one or more consequences are so unacceptable that safeguards are needed whatever the magnitude. Analogously, qualitative assessments can be used as a first step to quickly explore or implement protective measures where there is expert consensus that such measures would be immediately effective and useful. As such, if there are obvious sources of risk that can be eliminated, one does not need to wait for the results of a full quantitative risk assessment to implement risk management actions. A qualitative risk assessment may also provide the necessary insights into the pathway(s) associated with the risk of concern, but not previously identified, which also allows the risk manager to make decisions or apply safeguards without further quantification.

FAO/WHO (2004) noted:

> *"Qualitative risk assessments may be undertaken, for example, using the process of 'expert elicitation'. Synthesizing the knowledge of experts and describing some*

uncertainties permits at least a ranking of relative risks, or separation into risk categories. … As assessors understand how qualitative risk assessments are done, they may become effective tools for risk managers."

Noting that, in some circumstances, such as those indicated above, they can be conducted quickly and used to address specific questions and may reveal that an extensive, fully quantitative exposure, and risk assessment is not required to provide relevant advice to the risk manager.

9.1.2 *Qualitative risk assessment in food safety*

Qualitative risk assessments have been extensively used in import risk assessments of animals and of animal products (OIE, 2018). Many animal products are also food intended for human consumption; therefore, many of these import-risk assessments have also involved food products intended for human consumption. However, the focus of such import-risk assessments has historically been to assess the risk of a particular exotic pathogen entering a potential importing country or region, carried within the food in question. The intention is generally to assess whether the risk of importing the pathogen in the product is too high to be acceptable to the importing country, and whether safeguards should therefore be applied (such as cooking, freezing, testing or total ban). Frequently, further consequences, in particular any potential consequences to human health, have not been the focus of the risk assessment, even when the pathogen might be a zoonotic organism.

Food product import-risk assessments, in general, assess the probable presence of a pathogen in that product, so that if this probability is unacceptable, then import safeguards can be applied. Human health and safety risk assessments of food products, in general, not only set out to assess the probability of the presence of a pathogen, but also the amount of pathogen present, so that the human response to the probable dose can be assessed. The latter aspect is sometimes perceived to make qualitative risk assessments less useful in food safety applications, despite the fact that many quantitative dose-response data are very subjective in their estimation methods. However, not all steps in the risk assessment process (i.e. Hazard Identification, Hazard Characterization, Exposure Assessment, Risk Characterization) are necessary in all cases to assist food safety risk managers to deduce appropriate risk management actions. Actions to reduce exposure, even in the absence of dose-response data, would in many cases be appropriate risk management steps and could be determined from an 'incomplete' risk assessment (i.e. no Hazard Characterization), whether qualitative or quantitative. An epidemiologically based risk assessment may also not require dose-response data.

9.1.3 *Characteristics of a qualitative risk assessment*

The complementary nature of qualitative and quantitative risk assessments

The main principles of a risk assessment apply equally anywhere along the qualitative to quantitative risk assessment continuum. These include identification of the hazard, defining the risk question, outlining the steps of the risk pathway, gathering data and information, including information on uncertainty and variability, combining the information in a logical manner, and ensuring all is fully referenced and transparent. It follows from this that many of the activities are the same, up to and including the gathering of the data. Therefore, it is sometimes the case that a qualitative (or semi-quantitative) risk assessment is included in a risk profile, with the intention of following up with a quantitative risk assessment if it is subsequently thought to be necessary, or useful, and feasible.

The detailed investigative nature of a qualitative risk assessment may provide the risk manager or policymaker with all the information they require. A qualitative risk assessment may also provide the necessary insights into previously unidentified pathway(s) associated with the risk of concern, which allows the risk manager to make decisions or apply safeguards without further quantification. In

these circumstances additional quantitative assessments will probably be deemed unnecessary by the risk manager or policymaker.

A qualitative risk assessment can be informative even if a quantitative assessment is being planned. It can be used to identify the data currently available, the uncertainties surrounding that data, and uncertainties about exposure pathways, to decide if quantification is both feasible and likely to add anything to the current state of knowledge. It can identify areas of data deficiency for targeting future studies necessary prior to quantification. It can examine the probable magnitude of the risks associated with multiple risk pathways, such as exposure pathways, prioritizing them for the application of quantification.

Whatever the initial intention, when a qualitative risk assessment has already been undertaken, much of the work for a quantitative risk assessment has already been done. For the same risk question, quantification will be able to build on the risk pathway(s) and data already collected, to provide a numerical assessment of the risk.

Subjective nature of textual conclusions in qualitative risk assessments

Assessing the probability of any step in the risk pathway, or the overall risk, in terms of high, medium, low, negligible, etc., is subjective, as the risk assessor(s) will apply their own concepts of the meanings of these terms. These meanings may (and probably will) differ from person to person. This is one of the major criticisms levelled at qualitative risk assessments. However, these final risk assessors' estimates should never be viewed in isolation, just as numerical outputs from quantitative risk assessments should not, and reinforces the need for transparent documentation of the data and logic that lead to the assessor's estimate of the risk.

For a qualitative description of a risk to be useful to a risk manager, the assessor and manager must have similar perceptions of the meaning of subjective terms such as 'low', negligible', etc., or other descriptors (see also Section 7.2). A final risk characterization label, e.g. 'low', is largely meaningless to a risk manager without some sort of indication of what constitutes 'low' in the eyes of the author of the report. Also, it gives little indication of what particular pieces of evidence would change the assigned label to something other than 'low'. Thus, if evidence were to be presented that 25% of the product was not stored frozen, would the risk increase to moderate? Judgements will be used within any risk assessment. These may be the risk assessor's judgements, or expert opinion, or both, and these will always be subjective. This will apply when defining the scope of the problem, selecting (and rejecting) data, delineating the risk pathways, applying weightings to data or model pathways, selecting the distributions in a stochastic model, etc., as well as selecting a description of high, low, etc., in a qualitative assessment. Therefore, any risk manager, policymaker or other stakeholder who needs to use, or wishes to understand, a given risk assessment, irrespective of where on the qualitative to quantitative spectrum the risk assessment lies, should not simply look at the final 'result'. They should have some knowledge of how that result was arrived at.

Many people may not have the knowledge base to directly understand the computations involved within a quantitative risk assessment. They will need to rely on the explanations and opinions of the risk assessor in explaining to them how the result was reached, and what the underlying assumptions, judgements, uncertainties, etc., in the computation were. If the risk assessor is a good teacher as well as a good risk assessor, this can work well. But only under these circumstances are the risk managers likely to be able to decide for themselves the significance and meaning of the quantitative result.

As noted previously, the mathematical expression of risk inherent in a quantitative risk assessment may limit accessibility, unless accompanied by narrative explanations. Analogously, with a qualitative assessment, providing it has been written in a transparent and logical way, almost anyone should be able to understand and follow the arguments.

A definition of 'negligible' used in qualitative risk assessment is that, for all practical purposes, the magnitude of a negligible risk cannot, qualitatively, be differentiated from zero (for example, see the use of the term in OIE (2018). The term 'zero' is not used because in microbiological food safety there is generally no such thing as absolutely no risk. Note that, since 'negligible' may be understood as 'may be neglected', it can be argued to be a 'risk management' term because it involves a judgement.

It must be emphasised, that qualitative risk assessment relies on as much numerical data as possible to provide model inputs despite their textual nature, and the process of data gathering must be equally as thorough as for a quantitative risk assessment.

9.2 Semi-quantitative risk assessment

Semi-quantitative methods involve assigning labels to qualitative estimates in the form of probability ranges, weights or scores, and combining them by addition, multiplication, or other mathematical operation with the objective of achieving a greater level of objectivity compared to qualitative approaches. There must be a clear set of rules that dictate how the labels (scores, weights, ranges etc.) are combined. This set of rules should follow the basic probability principles, and be fully described and transparent regarding operation, and result generation. This provides an intermediary level between the textual evaluation of qualitative risk assessment and the numerical evaluation of quantitative risk assessment. It offers a more consistent and rigorous approach to assessing and comparing risks and risk management strategies than does qualitative risk assessment and avoids some of the greater ambiguities that a qualitative risk assessment may produce and it does not require the same mathematical skills as quantitative risk assessment. Semi-quantitative may be an attractive option when data are limited, but it should be noted that all forms of risk assessment require the greatest possible collection and evaluation of data available on the risk issue, and food safety risk assessments require in-depth knowledge in a variety of scientific disciplines. Semi-quantitative risk assessment requires all of the data collection and analysis activities for qualitative risk assessment as described in the previous section. It has been stated that semi-quantitative methods do not offer any advantages over well-researched, transparent, peer-reviewed qualitative approaches (OIE, 2018).

As noted in the previous section, Codex Alimentarius Commission (CAC) and others generally consider just two categories of risk assessment: qualitative and quantitative. Semi-quantitative risk assessment, as described here, has often been grouped together with qualitative risk assessment, but this understates the important differences between them in their structure and their relative levels of objectivity, transparency and repeatability.

9.2.1 *Uses of semi-quantitative risk assessment*

Semi-quantitative risk assessment is most useful in providing a structured way to rank risks according to their probability, severity or both (risk), and for ranking risk reduction actions for their effectiveness. This is achieved through a predefined scoring system that allows one to map a perceived risk into a category, where there is a logical and explicit hierarchy between categories.

Comparing hazards

One example of the utility of the semi-quantitative risk matrix approach is a probability-severity table. This approach offers quick way to visualize the relative riskiness (a term sometimes used for the combination of probability and severity) of several identified hazards within the domain of analysis. Table 34 illustrates an example, where all hazards (e.g. the list of pathogens that might appear in a particular food type) are recorded in one table, allowing for the easy identification of the most threatening hazards (those closer to the upper right corner) as well as providing a general picture of the overall risk associated with the food type. The numbers in the table are indices for identified hazards. Hazards 2 and 13, for example, have high risk; hazards 3 and 7 have very low risk. Hazards with zero events per year (i.e. zero probability; hazards 11 and 14) or no severity (hazards 8, 9 and 10) do not pose a risk, but may be useful to document as having been identified and subsequently determined to have negligible risk.

Table 34: Example of a P-I table for individual hazards (indicated by the numbers in the grid) per year (NIL=None, VLO = Very Low; Lo = Low; Med = Medium; Hi = High; VHI = Very High).

Severity							
	VHI			6			13,2
	HI	14				15	12
	MED		5		4	1	
	LO						
	VLO	11	7	3			
	NIL			8,9		10	
		NIL	**VLO**	**LO**	**MED**	**HI**	**VHI**
		Events per year					

Risk scores can then be used to rank the identified risks. A scaling factor, or score, is assigned to each label used to describe each type of severity. If a log scale is used to define each categorical scale, as in the example provided in Table 11 for probability, the probability and severity scores can be designed such that the risk score equals their sum, or some other simple mathematical equation. Table 35 provides an example of the type of scaling factors that could be associated with each probability and severity combination.

Table 35: Example risk score calculations for some hazards used in from Table 34.

Risk Index	Probability	Probability Score	Severity	Severity Score	Risk Score
13	VHI	5	VHI	6	5+6=11
1	HI	4	MED	3	4+3=7
5	VLO	1	MED	3	1+3=4

Comparing risks and risk management strategies

Semi-quantitative risk assessment is generally used where one is attempting to optimize the allocation of available resources to minimize the impact of a group of risks under the control of one organization. It helps achieve this in two ways: first the risks can be placed onto a sort of map so that the most important risks can be separated from the less important; second, by comparing the total score for all risks before and after any proposed risk reduction strategy (or combination of strategies) one can get a feel for how relatively effective the strategies are and whether they merit their costs.

Table 14 shows how a risk matrix might be separated into three regions. This is sometimes known as a 'traffic light' system: hazards lying in the green area are well within an acceptable level (low risk);

hazards lying in the red region are not acceptable (high risk); and the remaining hazards lie in the amber, or the medium risk, area. The crudeness of the scaling of this semi-quantitative risk assessment approach means that it will often be appropriate to study 'amber risks' further, perhaps using more quantitative methods, to determine whether they actually lie close to or within the red or green regions.

9.2.2 *Characteristics of a semi-quantitative risk assessment*

Categorical labelling is the basis for semi-quantitative risk assessment. It uses non-technical descriptions of a risk's probability, severity, and risk (the combination of probability and severity), for example: 'Very low', 'Low', Medium', 'High', and 'Very High', or some scaling like A-F. For this type of labelling to be unambiguous and useful, risk managers must provide a list of the non-overlapping, exhaustive categorical terms that are to be used, together with clear definitions of each term. For example, a 'Low' probability might be defined as an individual having between 10^{-3} and 10^{-4} probability of occurring in a year, and a 'High' severity might be defined as an individual suffering long-term sequelae that materially affect their quality of life. This step is crucial, as a number of studies have shown that even professionals, who are well-versed in probability ideas and who regularly make decision based on risk assessments, have no consistent interpretations of probability phrases ('likely', 'almost certain', etc.). This lack of consistent interpretation could lead to inconsistent assessment of risk and inadvertent lack of transparency. Without numerical definitions of probability, subjective descriptions such as 'low' can be affected by the severity: for example, a 5% probability of diarrhoeal illness from some exposure might be considered 'low', but a 5% probability of death from an exposure might be considered 'high'. The number of categories used to express probability and severity should be chosen so that one can be sufficiently specific without wasting time arguing about details that will not ultimately affect the risk management decision. A five-point scale has been the most commonly used in the risk community, sometimes with a sixth category representing zero for probability and severity, and a seventh 'certain' category for probability representing a probability of 1.

Often, in the course of carrying out a qualitative risk assessment, one can roughly estimate the probability of exposure, etc., from comparison with other, previously quantified risks or from good data pertaining to the problem in hand. If time or the available data are insufficient to carry out a complete quantitative risk assessment, one can use these categorical labels to express the risk level in a more structured way than a simple, qualitative description of the evidence one has acquired. An example is presented in Section 7.3.2.

9.2.3 *Limitations of semi-quantitative risk assessment*

A semi-quantitative risk assessment has its limitations and can cause errors in conclusions (see Cox Jr., 2008; Levine, 2012; Vatanpour, Hrudey and Dinu, 2015 for discussion on the issues with an emphasis on risk matrices). Issues arise from several difficulties in defining how categorical labels should be interpreted and manipulated. The risks are placed into usually quite broad sets of categories: it is common to use five or so for probability and for severity, not including zero, which gives 25 possible combinations. It is therefore imperative that the categories are carefully constructed. For example, one could break up the probability range into five categories, as in Table 36.

Table 36: A linear scoring system for probability.

Score	Probability range
1	0 – 0.2
2	0.2 – 0.4

3	0.4 – 0.6
4	0.6 – 0.8
5	0.8 – 1

However, under this scheme, a risk with a probability of 0.1 would sit in the same category (Score 1) as a risk with probability 0.000001, despite being 100 000 times more likely. This is one reason why a log scale is often chosen for probabilities. The nature of food safety risk means that often deals with probabilities that span several orders of magnitude, which also makes the use of a log scale more appealing and informative.

It cannot be easily combined probability scores for components of a risk pathway to get a probability score for the risks as a whole. For example, food safety risk estimation is often split into two parts: the probability of exposure; and the probability of illness given exposure. Using the scheme above, if the exposure had a 0.3 probability (score = 2) of occurring within a certain period for a random individual, and the probability of illness from that exposure was 0.7 (score = 4), the combined probability is 0.21 (0.3×0.7=0.21, which receives a score 2). It cannot be easily created a rule with scores that replicates the probability rules, and this limitation is well recognised (see references above). Taking the minimum of the two scores is one partial solution, but this generally over-estimates the risk. For example, changing the probability of illness given exposure to anything from 0.2 to 1.0 would give the same combined probability score of 2 using this approach.

The use of a log scale for probability relieves the problem to some extent if it is reversed the probability score order described so far to assign the highest probability with the lowest score, as shown in Table 37.

Table 37: A logarithmic scoring score order described so far to assign the highest system for probability.

Category	Probability range	Score
Negligible	Indistinguishable from 0	NA
Very Low	$< 10^{-4}$, (except 0)	5
Low	10^{-4} to 10^{-3}	4
Medium	10^{-3} to 10^{-2}	3
High	10^{-2} to 10^{-1}	2
Very High	$> 10^{-1}$ (except 1)	1
Certain	1	0

Using this scheme, the scoring system equivalent of multiplying probabilities is to add scores. For example, if the exposure had a 0.2 probability (score = 1) of occurring within a certain period for a random individual, and the probability of illness from that exposure was 0.004 (score = 3), the combined probability is 0.0008 (score 4). It does not always work out so neatly, however. An exposure with probability 0.5 (score = 1) and a probability of illness from that exposure of 0.003 (score = 3) gives a combined probability of 0.0015 (score = 3), yet the individual scores sum to 4. Adding scores in a log system like the one in Table 37 will often over-estimate the probability by one category. This is one reason for having an amber region in the traffic light system, because risks may be over-estimated, and risks falling into an amber region may in fact turn out to be acceptable.

There is also a problem of the granularity of the scale. For example, for a risk whose probability of occurrence falls just above the boundary between two categories, and for which a risk management strategy reduces that probability by a small amount, it could be dropped down one probability

category, which is now indistinguishable from reducing the probability by a factor of 10. However, there is nothing to stop the risk assessor from using score fractions if it seems appropriate. The integer system is designed for convenience and simplicity and could be changed to include fractions if this better represents the available knowledge.

Using the semi-quantitative risk assessment scoring system as a surrogate for probability calculations is also likely to cause more severe inaccuracies when one assesses a longer sequence of events. This is because the "errors" are being compounded; see for example the "Probabilities Are Inconsistent with Qualitative Aggregation Rules" (Cox Jr., Babayev and Huber, 2005)

10. Data

Risk assessment studies are developed by compiling information from a variety of data sources. Each of these data sources contributes in varying degrees to an understanding of the pathogen-host-matrix interactions (Figure 4) that influence the potential public health risks attributable to a disease agent. An appreciation of the strengths and limitations of the various data sources is critical to selecting appropriate data for use, and to establishing the uncertainty associated with different data sets and test protocols.

Active data collection is often required, because reliance on passive data submission or data in published form does not usually provide enough information in sufficient detail to construct the QRA model. Relevant data come preferably from peer-reviewed journals. In case of lack of data or insufficient data from published sources, it is also advisable to evaluate the availability of unpublished, high-quality data sources. Risk assessors should communicate with experimenters, epidemiologists, food or water safety regulators, and others who may have useful data that could contribute to the analysis. An example is the outbreak information collected by the Japanese Ministry of Health (Kasuga *et al.*, 2004) and which was used for dose-response modelling of *Salmonella*, along with other data (FAO/WHO, 2002a). When such data are used, the criteria and results of evaluation must be carefully documented. If using material published on the Internet, care should be taken to establish the provenance, validity and reliability of the data, and the original source, if known.

Understanding the characteristics of data sources is important to the selection and interpretation of data. Risk assessors often use data for a purpose other than that for which it was originally intended. Risk assessors and modelers need to know how the data they use were collected, and the purpose of their collection. The properties of the available data will depend on the perspective of the researchers generating the data (e.g. experimenter versus epidemiologist).

Two categories of data are necessary for the development of a risk assessment model: firstly, data that, in text format, describe the biological and physical processes as well as the human factors involved, and, secondly, numerical data that allow quantitative estimates to be calculated. The extent to which numerical data are required will vary from one risk assessment to another, depending on the defined purpose, scope, modelling approach and details chosen. In Table 38 an overview is presented of the types of data required for conducting a risk assessment as well as their collection sources and these are described in detail in the following sections.

3511 *Table 38: Data required for risk assessment and data collection sources*

Type of Data	Description	Collection Source
Hazard Identification		
Association between exposure and adverse health outcome	The evidence that can be utilized to pair the food and microbiological hazard and link the exposure to hazard in specific food to human illnesses	• Outbreak data • Foodborne disease surveillance and annual health statistics • Food safety rapid alert systems • Literature: Analytical epidemiological studies • Systematic food contamination monitoring surveys
Microbiological hazard characteristics	Characteristics of the organisms and mechanism with which the organism affects the host are described, while detailed dose-response analysis is done in hazard characterization	• Literature: microbiological studies
General characteristics of food and conditions of supply chain	Intrinsic characteristics of the food (e.g. pH, water activity) and process evaluation (e.g. time, temperature)	• Industry data and literature: Description of product and food supply
Adverse health outcomes in exposure population	Disease and sequelae in population and sub-populations by demographic and/or social-economic factors, sensitive population	• Scientific and medical literature
Hazard Characterization		
Parameters of dose response models	Parameters estimated by fitting to dose-response data to models	• Literature: dose-response fitted models
Dose response data	Data on dose response that can be fitted to a dose-response model	• Outbreak data • Volunteer feeding studies • Animal studies
Annual cases of the foodborne illness and prevalence of the pathogen in a food commodity	Data on reported cases of illness and prevalence of the causing pathogen in the food commodity to approximate a DR relationship	• Foodborne disease surveillance and annual health statistics • Systematic monitoring surveys
Exposure Assessment		
Prevalence and concentration	Data on prevalence and concentration of the pathogen in the food of concern at the starting point of the risk assessment and other points of the food chain	• Systematic food contamination monitoring surveys • Literature: prevalence and concentration surveys • Expert Knowledge Elicitation (EKE)

Type of Data	Description	Collection Source
Processing conditions	Data describing the conditions of food processing which may affect prevalence and concentration of the pathogen (i.e. time-temperature of thermal processing, fermentation, partitioning, etc)	• Literature: Description of product and supply chain • Industry data: Description of product and supply chain • EKE
Effect of processing stages and/or interventions	Data on the effect of a processing stage/intervention on prevalence and concentration of the pathogen	• Literature: Intervention studies • EKE
Product characteristics	Data on food characteristics (pH, aw, concentration of antimicrobials, packaging atmosphere, use-by-date, etc.) that may affect the behaviour of the pathogen during storage	• Literature: Description of product and supply chain • Industry data: Description of product and supply chain • EKE
Distribution and storage conditions	Time-temperature data for distribution and storage of the food at retail and domestic level	• Literature • Industry data: Description of product and supply chain • EKE
Conditions of food handling and preparation	Data describing the conditions of food handling and preparation which may affect prevalence and concentration of the pathogen (i.e. time-temperature of cooking, partitioning, etc.)	• Literature: cross contamination, food handling and preparation • EKE
Kinetics of pathogen's behaviour	Data on the kinetics of pathogen's growth/survival/inactivation during food processing, distribution, storage, handling and cooking.	• Literature: predictive microbiology models • Modelling online tools
Consumption	Data on serving size Data on frequency of consumption, and number of annual servings for different population groups (normal, susceptible, pregnant, etc.).	• National consumption databases • EKE • Total diet studies
Population segments	Data on population size by segments	• National population census
Annual production of the food commodity	Data on tons of food produced in a country and information of imports, if necessary	• National food production statistics
Risk Characterization		
Annual cases of the foodborne illness	Data used for anchoring and/or validating a risk assessment model	• Foodborne disease surveillance and annual health statistics

3512

Data should be collected to represent reality as closely as possible. The same principles can be applied, for instance, to fisheries as to primary production, or to food service (catering) as the point of consumption, as well as for issues related to waterborne microbiological hazards. Note that the specific scope and purpose of a risk assessment can be much narrower in practice and these will determine the type and detail of data required. Since data are not available in all instances, alternative (surrogate) data may be employed. It is important to clearly describe the rationale and suitability for selecting the alternative data and evaluate the influence of using such data on the final risk estimates (Chapters 14 and 15)..

This chapter presents a summary of the types of data typically required for the construction of a risk assessment, capturing in brief the strengths and limitations of each of several classes of data sources.

10.1 Literature (primary and/or meta-analysis)

Data required for risk assessments may come from a wide variety of published sources, some of which may be common to many countries. Academia and other organizations publish their findings in the public domain. This can be in the form of documents that have been peer-reviewed within the scientific community or via non-peer reviewed written communications (conference proceedings, books, internet sites). Data from different sources may be helpful in confirming the degree of scientific agreement or uncertainty on a particular point.

In most cases, data need to be extracted from sources that are not intended for that specific purpose. Consequently, data may not be readily available in the exact form or detail required for the risk assessment. At this point, meta-analysis can be considered as a useful tool for combining or pooling the data from different sources in a structured way. In building up the risk assessment model, separate meta-analyses can be carried out to model the overall effect of certain processing stage or intervention strategy on likelihood/concentrations of a pathogen, as well as likelihood/concentrations of a pathogen at a starting point of the model or at any particular point in the food chain (Gonzales-Barron *et al.*, 2017). Multilevel meta-analysis models that account for the effect of selected *moderators* can also be used. For example, in Prado-Silva *et al.* (2015), such models were developed to summarize the effects of sanitizing treatments on *Salmonella* spp., *Escherichia coli* O157:H7 and *Listeria monocytogenes* in fresh produce, as affected by type of sanitizer and washing time and temperature.

Risk assessors familiar with meta-analysis techniques may conduct meta-analysis on summary statistics or on raw data. Original data may need to be requested from authors when the data are critical for a risk assessment. Human resources and availability of sufficient primary research sources will constrain the use of meta-analysis.

Scientific publications often give a good level of detail about the subject matter being investigated. The conditions under which the data were obtained, and the methods used are often well documented. If a number of individual studies addressing the same research question have been found, meta-analysis can be conducted to obtain a more reliable and representative overall estimate.

However, a drawback of published research is that, in many cases, aggregate data rather than raw data are published and that raw data may be difficult to access. Some journals are encouraging authors to make their raw data available, e.g. International Journal of Food Microbiology. The diversity in languages used for publications can pose a barrier to general access and use. Uncertainty and variability in the data are generally not described, and authors might need to be consulted to

obtain information on those aspects. Some research may be published but hard to locate due to a lack of readily accessible computer listings for items like fact sheets, conference proceedings, theses, dissertations, etc.

Another potential downside of published research is the potential for publication bias. This type of bias occurs because publishers prefer to publish novel research findings, rather than confirmatory research. As a result, the reported effects, e.g. for the efficacy of an intervention, may be larger than what might be expected in general and this type of 'error' has been referred to as a *Magnitude Error (M-Error)* by Gelman and Carlin (2014). Publication bias directly affects meta-analysis, although there are procedures to adjust the meta-analytical estimates when publication bias is likely to be present (Rothstein, Sutton and Borenstein, 2006).

10.1.1 *Analytical epidemiological studies*

Epidemiological surveys concern studies that have been commissioned to specifically investigate the causal relationship between the occurrence of foodborne illness and exposure to certain microbiological hazards through food consumption. They are most commonly undertaken as part of outbreak investigations, e.g. case-control or cohort study. These studies can be useful for hazard identification and characterization.

Strength

Epidemiological studies are very specific and provide a large amount of detailed information on the hazard and the consumer group investigated.

Limitations

Data are often generated for a relatively small number of consumers, and thus are not representative of larger consumer groups.

10.1.2 *Microbiological studies of prevalence and counts/concentrations*

The microbiological studies discussed here refer to studies reporting the prevalence and count/concentration of target microorganisms at various stages and those studies reporting the change in these, such as the efficacy of a processing intervention. These studies may report findings throughout the production and processing chain, including in the final food product. They are especially useful for the exposure assessment but may also inform the hazard identification.

Strength

Results from those studies provide useful information as the initial input or data needed for connecting the parts of the exposure model (e.g. Figure 5). Microbiological surveys undertaken at retail can provide valuable data to verify that the exposure model prediction (up to the retail stage) are comparable with what is observed at retail, i.e. a reality check.

For studies related to interventions (or growth or survival) the processing conditions such as durations, temperatures, etc are often reported and these provide useful inputs into predictive microbiological models.

Limitations

These studies often present results in an aggregate form, e.g. mean and standard deviation. Where possible, the raw data, without identifying information, should be requested from the authors as this will allow more detailed interrogation of the data than may be presented in a scientific publication. This allows statistical distributions of the data to be better ascertained and summary statistics (including variances) to be calculated and assessed for different components of the study. Such

intricacies may not be included in the scientific publication, possibly because the data were not specifically collected for use in a risk assessment.

It often happens that different laboratories use different microbial testing methods that are not measuring the same feature. Therefore, when reviewing published articles investigating the same research question, comparability of results should be appraised to see if these sources are in effect measuring the same thing or not, and if so, whether the same level of uncertainty exists. Differences in testing method comparability are probably the most difficult to resolve when attempting to compare final estimates. Ensuring that internationally validated microbiological methods are used can facilitate this comparison. For example, in some tests, different laboratories may use methods with different detection limits. Nevertheless, there have been advances to take into account the limit of detection and/or quantification when analysing data, without having to resort to biased 'substitution' methods (Shorten, Pleasants and Soboleva, 2006; Lorimer and Kiermeier, 2007; Gonzales-Barron *et al.*, 2010; Busschaert *et al.*, 2011; Williams and Ebel, 2014).

10.1.3 *Cross-contamination data during food processing*

The potential for microbial cross-contamination within the food processing environment is well recognized. Data and models that give insight into the extent to which this occurs is therefore required. Important areas will include, for instance, the level of contact between live and slaughtered animals or between raw and processed vegetable material, worker hygiene, operating equipment, plant design, sanitation protocols, and methods of packaging (Gallagher *et al.*, 2016; Pouillot *et al.*, 2015a; e.g. Zoellner *et al.*, 2019).

Strength

These studies can provide quantitative information on the frequency, extent and type of cross-contamination event that occur in a food processing environment. This allows better modelling of the cross-contamination as part of exposure assessment.

Limitations

Due to the amount of time involved in observing a reasonable number of cross-contamination events, and the variability between observation times (e.g. days or shifts), these types of studies likely involve only one or, a few different, food processing environments. Consequently, the results may be specific to the environment that has been observed and may not be representative of the industry as a whole.

10.1.4 *Food handling and preparation*

Storage and preparation practices, both in the home and in the catering environment, can influence the level of exposure. In particular, hazard growth or reduction may occur during storage prior to preparation if the temperature favours either of these processes; reduction in hazard contamination may occur as a result of cooking; and hazard concentration in cooked products may increase due to cross-contamination. To address these issues, data should be accompanied by descriptions of relevant details such as: times and temperatures of storage; typical handling practices and the potential cross-contamination events that could occur during preparation; the extent to which these events occur and the likely numbers of organisms transferred to different locations within the kitchen; the extent to which consumers are exposed to the organisms that are transferred; and typical cooking times and temperatures. Predictive microbiology models will be needed for these stages as well to assess potential changes in levels of pathogens and the resultant effect on risk.

Research has been undertaken on consumer practices, although the work tends to be product and situation specific (e.g. DeDonder *et al.*, 2009; Kosa *et al.*, 2015). As a result, still relatively little

information exists on food handling practices in the home that affect the safety of foods, although this situation is gradually changing (Murray *et al.*, 2017; Young *et al.*, 2017a). Food handling practices vary by geographical region or even within the same country, based, for example, on ethnicity, gender and education. Consumer storage times, extent of cross-contamination, cooking times and temperatures (such as reported by EcoSure, 2008), hot holding temperatures and times, and other data are not generally available. Likewise, relatively little information is available about food handling practices by restaurant and food service operations, including street food, which accounts for an increasingly greater proportion of meals in many countries, and this data gap is also gradually being addressed (Pichler *et al.*, 2014; Samapundo *et al.*, 2015; Tessema, Gelaye and Chercos, 2014). Some research is now being undertaken using human volunteers who are asked to prepare specific foods in custom kitchen that allow observation and video recording of the study participants, so that food handling practices can be quantified and objectively evaluated against prespecified criteria.

Strength

Directly observing food handling practices and measuring food storage, cooking or associated metrics (e.g. temperature) are more reliable than information obtained through an interview. That is, observation allows recording of what people do, rather than what they say they do.

Recording video footage of food preparation is also a good way to reduce researcher bias. The actual food handling practices can be assessed "blindly" through an independent third-party. However, care must be taken that the specific practices that are assessed have been well described and documented to ensure consistency.

Limitations

It is difficult to observe food handling practices directly as they are practiced in homes and food service operations, especially when researchers want to capture video footage of the food handling. The best alternative is to use purpose-built food preparation kitchens that allow observation. However, these are costly to establish and to maintain (including the qualified staff to undertake studies).

These types of studies can pose ethical problems and because they cannot be undertaken in a 'blind' way, i.e. volunteers know that they are being observed, may change the way the food is handled.

Where measurements are involved (e.g. EcoSure, 2008) care must be taken that equipment is properly calibrated and that raw data are critically checked for recording errors.

10.1.5 *Human volunteer feeding studies*

The most obvious means for acquiring information on dose-response relations for foodborne and waterborne pathogenic microorganisms is to expose humans to the disease agent under controlled conditions. There have been a limited number of pathogens for which feeding studies using volunteers have been carried out. Most have been in conjunction with vaccine trials. Examples of the use of volunteer studies to develop dose-response models for a range of enteric pathogens are provide by Teunis *et al.* (1996), which includes references to the original experimental studies.

These studies are generally conducted only with primarily healthy individuals between the ages of 18 and 50, and thus do not examine the segments of the human population typically most at risk. Pathogens that are life threatening or that cause disease only in high-risk subpopulations are not amenable to volunteer studies. Typically, the studies investigate a limited number of doses with a limited number of volunteers per dose. The dose ranges are generally high to ensure a response in a

significant portion of the test population, i.e. the doses are generally not in the region of most interest to risk assessors.

The process of (self-)selection of volunteers may induce bias that can affect interpretation of findings. Feeding studies are not a practical means to address strain virulence variation. The choice of strain is therefore a critical variable in such studies. Most feeding studies use only rudimentary immunological testing prior to exposure. More extensive testing could be useful in developing susceptibility biomarkers.

Usually, feeding studies involve only a few strains, which are often laboratory domesticated or collection strains and may not represent wild-type strains. In addition, the conditions of preparation immediately before administration are not usually standardized or reported, though these may affect tolerance to acid, heat or drying, as well as altering virulence. For example, passage of *Vibrio cholerae* through the gastrointestinal tract induces a hyper-infectious state, which is perpetuated even after purging into natural aquatic reservoirs. This phenotype is expressed transiently, and lost after growth in vitro (Merrell *et al.*, 2002). In many trials with enteric organisms, they are administered orally with a buffering substance, specifically to neutralize the effect of gastric acidity, which does not directly translate into what the dose response would be if ingested in food or water.

Strengths

Using human volunteers is the most direct means of acquiring data that relates an exposure to a microbial hazard with an adverse response in human populations. If planned effectively, such studies can be conducted in conjunction with other clinical trials, such as the testing of vaccines. The results of the trials provide a direct means of observing the effects of the challenge dose on the integrated host defence response. The delivery matrix and the pathogen strain can be varied to evaluate food matrix and pathogen virulence effects.

These studies can provide information on both infection, e.g. by testing faecal matter for the hazard of interest, and illness, e.g. by observing symptoms in the volunteers.

Limitations

There are severe ethical and economic limitations associated with the use of human volunteers; especially because of the ethical implications these studies are no longer undertaken. However, for the purpose of better interpretation and utilization of the data reported in the literature, the aspects that are commonly considered in the development and assessment of an experimental design are listed below.

- What isolate, species, serotype and/or genotype, strain, etc. of the hazard was used?
- How is dose measured (both units of measurement and the process used to measure a dose)?
- How do the units in which a dose is measured compare with the units of measurement for the hazard in an environmental sample?
- Total units measured in a dose may not all be viable units or infectious units.
- Volunteers given repeat doses may not all receive the same amount of inoculum.
- How is the inoculum administered? Does the protocol involve simultaneous addition of agents that alter gastric acidity or promote the passage of microorganisms through the stomach without exposure to gastric acid?

- How is it known that the volunteers are naïve (serum antibodies may have dropped to undetectable levels or the volunteer may have been previously infected with a similar pathogen that may not be detected by your serological test)?
- How is infection defined?
- What is the sensitivity and specificity of the assay used to determine infection?
- How is illness defined?

10.1.6 *Animal studies*

Animal studies are used to overcome some of the logistical and ethical limitations that are associated with human-volunteer feeding studies. There are a large variety of different animal models that are used extensively to understand the hazard, host and matrix factors that affect characteristics of foodborne and waterborne disease, including the establishment of dose-response relations.

Strengths

The use of surrogate animals to characterize microbial hazards and establish dose-response relations provides a means for eliminating a number of the limitations of human-volunteer studies while still maintaining the use of intact animals to examine disease processes. Animal models can be relatively inexpensive, thus increasing the potential for testing a variety of strains and increasing the number of replicates and doses. The animals are generally maintained under much more controlled conditions than human subjects. Immunodeficient animal strains and techniques for suppressing the immune system and other host defences are available and provide a means for characterizing the response in special subpopulations. Testing can be conducted directly on animal subpopulations such as neonates, aged or pregnant populations. Different food vehicles can be investigated readily.

Limitations

A major limitation is that the response in the animal model has to be extrapolated to that in humans. There is seldom a direct relationship between the response in humans and that in animals. Often, differences between the anatomy and physiology of humans and animal species lead to substantial differences in dose-response relations and the animal's response to disease. For a number of food pathogens, it can be challenging to select an appropriate animal model, as the successful extrapolation from the animal to the human population depends on several factors, such as the similarity of pathogenic mechanisms, the physiological and immune responses between animals and humans (Buchanan, Smith and Long, 2000). Several highly effective models (e.g. primates or pigs) can be expensive and may be limited in the number of animals that can be used per dose group; ethical concerns over animal experimentation need to be carefully considered. Some animals used as surrogates are highly inbred and consequently lack genetic diversity. Likewise, they are healthy and usually of a specific age and weight range. As such, they generally do not reflect the general population of animals of that species, let alone the human population. Ethical considerations in many countries limit the range of biological end-points that can be studied.

When human-derived data are absent, the validation of dose-response models built upon animal studies is challenging. However, there are some general considerations regarding animal models to narrow the difference between animal models and human target. When surrogate pathogens or surrogate animal models are used, the biological basis for the use of the surrogate must be clear. Using data obtained with animal models to predict health effects in humans could take advantage of the use of appropriate biomarkers. It is important to use pathogen strains that are identical or closely related to the strain of concern for humans, because, even within the same species and

subspecies, different strains of pathogens may have different characteristics that cause variation in their abilities to enter and infect the host and cause illness.

10.1.7 *In-vitro studies*

In vitro studies involve the use of cell, tissue or organ cultures and related biological samples to characterize the effect of the hazard on the host. They are of most use for qualitative investigations of pathogen virulence but may also be used to evaluate in detail the effects of defined factors on the disease process. For example, the effect of food processing and preservation conditions on a pathogen's virulence and toxin production can be evaluated by *in vitro* studies (Greppi and Rantsiou, 2016; Haddad *et al.*, 2018).

Strengths

In vitro techniques can readily relate the characteristics of a biological response with specific virulence factors (genetic markers, surface characteristics and growth potential) under controlled conditions. This includes the use of different host cells or tissue cultures to represent different population groups, and manipulation of the environment under which the host cells or tissues are exposed to the pathogen, to characterize differences in dose-response relations between general and special populations. In vitro techniques can be used to investigate the relations between matrix effects and the expression of virulence markers. Large numbers of replicates and doses can be studied under highly controlled conditions.

These techniques can be used to readily compare multiple species and cell types to validate relationships between humans and surrogate animals. They are particularly useful as a means of providing information concerning the mechanistic basis for dose-response relations.

Limitations

The primary limitation is the indirect nature of information concerning dose-response relations. One cannot directly relate the effects observed with isolated cells and tissues to disease conditions that are observed within intact humans, such as the effect of integrated host defences. To compare with humans, there is need for a means to relate the quantitative relations observed in the in vitro system to those observed in the host. For many organisms, the specific virulence mechanisms and markers involved are unknown, and may vary between strains of the same species.

Similar to some other data types, such as public health surveillance, these types of studies are usually limited to providing details of factors affecting dose-response relations and to augmenting the hazard characterization but are unlikely to be a direct means of establishing dose-response models useful for risk assessments.

10.1.8 *Biomarkers*

Biomarkers are measurements of host characteristics that indicate exposure of a population to a hazard or the extent of adverse effect caused by the hazard. Examples include serological assays, counts of subsets of white blood cells and production of gaseous oxides of nitrogen. Biomarkers are generally minimally invasive techniques that have been developed to assess the status of the host. Also 'omics' (transcriptomics, metabolomics) type biomarkers can be used (Haddad *et al.*, 2018). The United States National Academy of Science has classified biomarkers into three classes (NRC, 1989; Slikker Jr., 2018), as follows:

- Biomarker of exposure – an exogenous substance or its metabolite, or the product of an interaction between a xenobiotic agent and some target molecule or cell, that is measured in a compartment within an organism.

- Biomarker of effect – a measurable biochemical, physiological or other alteration within an organism that, depending on magnitude, can be recognized as an established or potential health impairment or disease.
- Biomarker of susceptibility – an indicator of an inherent or acquired limitation of an organism's ability to respond to the challenge of exposure to a specific xenobiotic substance.

Even though this classification was developed against the background of risk assessment of toxic chemicals, these principles can be useful in interpreting data on microbial hazards. In future also the gut microbiome might be related to disease susceptibility.

Strengths

These techniques provide a means of acquiring biologically meaningful data while minimizing some of the limitations associated with various techniques involving human studies. Typically, biomarkers are measures that can be acquired with minimum invasiveness while simultaneously providing a quantitative measure of a response that has been linked to the disease state. As such, they have the potential to increase the number of replicates or doses that can be considered, or to provide a means by which objectivity can be improved, and increased precision and reproducibility of epidemiological or clinical data can be achieved. Biomarkers may also provide a means for understanding the underlying factors used in hazard characterization. A biomarker response may be observed after exposure to doses that do not necessarily cause illness (or infection). Lefkowitz *et al.* (1992) noted antibodies to *Vibrio vulnificus* in shellfish industry workers. Biomarkers can be used either to identify susceptible populations or to evaluate the differential response in different population subgroups. Egorov *et al.* (2018) noted the application of salivary immunoassay in a prospective community study of waterborne infections.

It should also be noted that the most useful biomarkers are linked to illness by a defined mechanism, that is, the biological response has a relationship to the disease process or clinical symptom. If a biomarker is known to correlate with illness or exposure, then this information may be useful in measuring dose-response relationships, even if the subjects do not develop clinical symptoms. Biomarkers such as these can be used to link animal studies with human studies for the purposes of dose-response modelling. This is potentially useful because animal models may not produce clinical symptoms similar to humans. In which case, a biomarker may serve as a surrogate end-point in the animal.

Limitations

Biomarkers are often indicators of infection, illness, severity, duration, etc. As such, there is a need to establish a correlation between the amplitude of the biomarker response and illness conditions. Biomarkers primarily provide information on the host status, unless protocols are specifically designed to assess the effects of different pathogen isolates or matrices.

The only currently available biomarkers for foodborne and waterborne pathogens are serological and salivary assays. The main limitation for such assays is that, in general, the humoral immune response to bacterial and parasitic infections is limited, transient and non-specific. For example, efforts to develop an immunological assay for *Escherichia coli* O157 infections have shown that a distinctive serological response to the O antigen is seen typically in the most severe cases, such as those with bloody diarrhoea, but can be absent in less severe cases, such as cases with blood-less diarrhoea. In contrast, serological assays are often quite good for viruses.

Another limitation is that some biomarkers, such as serological assays, can result in false positives. For serological assays, the presence of antibodies that cross-react with microbial antigens used in

the assay or interfering substances that interact with assay components can also lead to false-positive results. Thus, positive Immunoglobulin M (IgM) assay results require cautious interpretation – consideration of clinical course compatibility and epidemiological factors – and/or confirmation by other serological or molecular testing methods (Woods, 2013).

Other biomarkers, such as counts of subsets of white blood cells or production of gaseous oxides of nitrogen are possible but have not been tested extensively in human populations.

10.2 National and international surveillance data

10.2.1 *Food safety rapid alert systems*

A food safety rapid alert system allows national food and feed control authorities to share information about measures taken in response to serious risks detected in relation to food, and as such can provide useful information for hazard identification. This exchange of information helps countries to act more rapidly and in a co-ordinated manner in response to health threats caused by food. One example of a food safety rapid alert system is the *European Rapid Alert System for Food and Feed*[11] (RASFF). Through the RASFF consumers' portal, the latest information on food recalls, public health warnings and border rejections in all EU countries can be accessed.

The functioning principle of the RASFF is simple: if a member of the network has any information relating to the existence of a direct or indirect risk to human health deriving from food or feed, that information must be immediately notified to the Commission and, where the EFTA States are involved, to the Authority. The Commission disseminates this information immediately to all members of the network.

Strengths

This system enables data sharing between geographically linked parties in an efficient manner. The data should be representative of the food within a diverse but geographically linked region.

Limitations

The system is only as good as it's least active member. If one country does not have the resources or expertise to easily contribute data, then the resulting dataset is limited or skewed toward the other countries in the system.

Similarly, the system will likely have good information about common and well recognized hazards, which tend to be part of national surveillance activities or outbreak investigations (Sections 10.2.1 and 10.2.3). Emerging hazards, those that are not actively surveyed or those that do not require reporting under a national health system may be less likely to be captured in a rapid alert system, unless a large enough outbreak has been identified and reported.

While rapid alert systems can be excellent sources of information for when a hazard has been identified in a food, they usually do not provide useful information about prevalence of the hazard occurring. This is because to the denominator is not generally captured and information about the food units in which a hazard has not been detected are not reported. In addition, if a hazard has not been reported for a particular food product in a rapid alert system, then this does not imply that the hazard does not occur in that food – it simply means that the hazard:food combination has not be

[11] https://ec.europa.eu/food/safety/rasff_en accessed 20 June 2019

reported in the system, either because the food has not been tested for the hazard, or because the hazard has not (yet) been detected in the food.

10.2.2 *Outbreak data*

When there is a common-source outbreak of foodborne or waterborne disease of sufficient magnitude, an epidemiological investigation is generally undertaken to identify the cause of the problem, to limit its further spread, and to provide recommendations on how the problem can be prevented in the future. Such information can be particularly valuable for hazard identification and characterization.

An outbreak of confirmed aetiology that affects a clearly defined group can provide very good information about the range of illness that a pathogen can cause, particular behaviour or other host characteristics that may increase or decrease the risk, and – if there is clinical follow up – the risk of sequelae. When the outbreak is traced to a food or water source that can be quantitatively cultured under circumstances that allow the original dose to be estimated, the actual dose-response can be estimated. Even when that is not possible, dose-effect relations can often be observed that show variation in clinical response to changes in relative dose and is part of the classic approach to an outbreak investigation. This may include looking for higher attack rates among persons who consumed more of the implicated vehicle but may also include variation in symptom prevalence and complications. There are good public health reasons for gathering information on the amount of the implicated food or water consumed. An outbreak that is characterized by a low attack rate in a very large population may be an opportunity to define the host-response to very low doses of a pathogen, if the actual level of contamination in the food can be measured. In addition, data from outbreaks are the ultimate "anchor" for dose-response models and are an important way to validate risk assessments (see also Section 16.2.2).

In general, information on several outbreaks– including the dose and the attack rate – is needed to establish a dose-response model, as each outbreak essentially contributes one data point to which the dose-response model is fitted. Examples include the dose-response models for *Salmonella* (FAO/WHO, 2002a) and *E. coli* O157:H7 (Strachan *et al.*, 2005).

Strengths

An outbreak investigation can capture the diversity of host response to a single pathogenic strain, down to the DNA level, e.g. using whole genome sequencing (e.g. Smith *et al.*, 2019). This can include the definition of the full clinical spectrum of illness and infection, if a cohort of exposed individuals can be examined and tested for evidence of infection and illness, e.g. using a case-control study. This may be undertaken independent of whether they were ill enough to seek medical care or diagnose themselves. It also includes definition of subgroups at higher risk, and the behaviour, or other host factors, that may increase or decrease that risk, given a specific exposure. Collecting information on underlying illness or pre-existing treatments is routine in many outbreak investigations.

Obtaining highly specific details of the food source and its preparation in the outbreak setting is often possible, because of the focus on a single food or meal, and may suggest specific correlates of risk that cannot be determined in the routine evaluation of a single case. Often, the observations made in outbreaks suggest further specific applied research to determine the behaviour of the pathogen in that specific matrix, handled in a specific way. For example, after a large outbreak of shigellosis was traced to chopped parsley, it was determined that *Shigella sonnei* grows abundantly on parsley left at room temperature if the parsley is chopped, but does not multiply if the parsley is

intact (Wu *et al.*, 2000). Such observations are obviously important to someone modelling the significance of low-level contamination of parsley.

Where samples of the implicated food or water vehicle can be quantitatively assayed for the pathogen, in circumstances that allow estimation of the original dose, an outbreak investigation has been a useful way to determine the symptoms associated with a defined dose in the general population.

Follow-up investigations of a (large) cohort of cases identified in an outbreak may allow identification and quantification of the frequency of sequelae, and the association of sequelae with specific strains or subtypes of a pathogen.

If preparations have been made in advance, the outbreak may offer a setting for the evaluation of methods to diagnose infection, assess exposure or treat the infection.

Limitations

The primary limitation is that the purpose and focus of outbreak investigations is to identify the source of the infection to prevent additional cases, rather than to collect a wide range of information. The case definitions and methods of the investigation are chosen for efficiency, and often do not include data that would be most useful in a hazard characterization and may vary widely among different investigations. The primary goal of the investigation is to quickly identify the specific source(s) of infection, rather than to precisely quantify the magnitude of that risk. Key information that would allow data collected in an investigation to be useful for risk assessments is therefore often missing or incomplete. Estimates of dose or exposure in outbreaks may be inaccurate because:

- It was not possible to obtain representative samples of the contaminated food or water.
- If samples were obtained, they may have been held or handled in such a way, after exposure occurred, as to make the results of testing meaningless. For example, microbial growth may have occurred if food is held at room temperature for extended periods.
- Laboratories involved in outbreak testing are mainly concerned with presence/absence, and they may not be conducting enumeration testing.
- It is very difficult to detect and quantify viable organisms in the contaminated food or water, e.g. viable *Cryptosporidium* oocysts in water or norovirus in oysters.
- Estimates of amount consumed by infected (and not infected) individuals, and of the variability therein, are poor.
- There is inadequate knowledge concerning the health status of the exposed population, and the number of individuals who consumed food but did not become ill (a part of whom may have developed asymptomatic infection, whereas others were not infected at all).
- The size of the total exposed population is uncertain.

In such instances, use of outbreak data to develop dose-response models generally requires assumptions concerning the missing information. Fairly elaborate models may be necessary to reconstruct exposure under the conditions of the outbreak. If microbiological risk assessors and epidemiologists work together to develop more comprehensive outbreak investigation protocols, this should promote the collection of more pertinent information. This might also help to identify detailed information that was obtained during the outbreak investigation but was not reported.

Even when all needed information is available, the use of such data may bias the hazard characterization if there are differences in the characteristics of hazard strains associated with

outbreaks versus sporadic cases, see for example Frank *et al.* (2014). The potential for such bias may be evaluated by more detailed microbiological studies on the distribution of growth, survival and virulence characteristics in outbreak and endemic strains.

Attack rates may be overestimated when they are based on signs and symptoms rather than laboratory-confirmed cases. Alternatively, in a case-control study conducted to identify a specific food or water exposure in a general population, the attack rate may be difficult to estimate, and may be underestimated, depending on the thoroughness of case finding.

The reported findings depend strongly on the case-definition used. Case definitions may be based on proximity in time and geography, clinical symptoms, on laboratory data or a combination thereof. The most efficient approach could be to choose a clinical case definition and validate it with a sample of cases that are confirmed by laboratory tests. This may include some non-specific illnesses among the cases. In investigations that are limited to culture-confirmed cases, or cases infected with a specific subtype of the pathogen, investigators may miss many of the milder or non-diagnosed illness occurrences, and thus underestimate the risk. The purpose of the outbreak investigation may lead the investigators to choices that are not necessarily the best for hazard characterization.

While outbreaks can be a valuable source of information for hazard identification and characterization, an outbreak ultimately only provides one data point – a combination of dose and estimated proportion of infected or ill persons. Consequently, numerous outbreaks involving the hazard (though possibly different strains) are required to allow the fitting of a dose-response model.

10.2.3 *Foodborne disease surveillance and annual health statistics*

Countries and several international organizations compile health statistics for infectious/zoonotic diseases, including those that are transmitted by foods and water. The data included in many cases are very specific, with rather detailed descriptions of the food (e.g. type, amount, composition), pathogen (reliably identified, often subtyped) and consumer (e.g. age, gender, health condition) being collected, often in the pursuit of identifying and investigating outbreaks (see also Section 10.2.1). Enhanced surveillance networks have in recent years improved the accumulation of data generated in foodborne disease investigations. These include Foodnet[12], Pulsenet[13] and Pulsenet International[14]. Such data are critical to adequately identify and characterize microbial hazards in specific food products.

In cases where no surveillance data or health statistics are available, it may be possible to use surrogate sources, if they are available. For example, for infections involving *Taenia saginata* sales data of taenicidal drugs have been used as an indication of the public health burden (Dorny and Praet, 2007).

Strengths

Active public health surveillance for foodborne illness can provide useful information about different disease endpoints and their proportional likelihood. Depending on the amount of information available different estimates may be obtained for various subpopulations of interest. However, care must be taken to account for the effect of under-reporting, which, depending on the hazard, may be substantial (e.g. Hall *et al.*, 2008; Scallan *et al.*, 2011).

[12] http://www.cdc.gov/foodnet/ accessed 20 June 2019
[13] http://www.cdc.gov/pulsenet/ accessed 20 June 2019
[14] http://www.pulsenetinternational.org/ accessed 20 June 2019

Annual health statistics provide one means of both anchoring and validating dose-response models (see Sections 16.2.2 and 16.2.3). The effectiveness of dose-response models is typically assessed by combining them with exposure estimates and determining if they approximate the annual disease statistics for the hazard; this process is sometimes referred to as a 'reality check.'

In addition, surveillance statistics may provide useful information about different morbidity ratios, i.e. rates with which different severities are observed. For example, Scallan *et al.* (2011) provide information on hospitalization and mortality rates and these differ between hazards or between different countries/regions (WHO, 2015). Similarly, surveillance information together with microbiological and genomic analyses can support the understanding of the severity of a hazard, e.g. some STEC strains have greater potential to cause more severe illness (from diarrhoea, bloody diarrhoea to haemolytic uremic syndrome) illness than others (FAO/WHO, 2018a).

In addition, annual disease statistics data have been used in conjunction with food survey data to rapidly estimate a simple dose-response relationship. It must be noted that, usually, analysis of such aggregated data requires many assumptions to be made, which increases the uncertainty in the results. This approach is highly cost–effective since the data are generated and compiled for other purposes. Available databases often have sufficient detail to allow consideration of special subpopulations.

Limitations

The primary limitations of these data are that they are highly dependent on the adequacy and sophistication of the surveillance system used to collect the information; and data only concern a limited range of microbiological hazards and do not necessarily reflect sporadic cases. Typically, public health surveillance for foodborne diseases depends on laboratory diagnosis. Thus, it only captures those who were ill enough to seek care (and were able to pay for it), and who provided samples for laboratory analysis. This can lead to a bias in hazard characterizations toward health consequences associated with developed nations that have an extensive disease surveillance infrastructure. Within developed countries, the bias may be towards diseases with relatively high severity, that more frequently lead to medical diagnoses than mild, self-limiting diseases. Comparisons with other countries are difficult because a set of defined criteria for reporting is lacking at an international level.

Another major limitation in the use of surveillance data is that it seldom includes accurate information on the attribution of disease to different food products, on the levels of disease agent in food and the number of individuals exposed. Use of such data to develop dose-response relations is also dependent on the adequacy of the exposure assessment, the identification of the portions of the population actually consuming the food or water, and the estimate of the segment of the population at increased risk. Nevertheless, these national surveillance data have been used in combination with national production or consumption data to provide crude comparisons across commodities (Hsi *et al.*, 2015).

10.2.4 *Systematic food contamination monitoring surveys*

Frequently, governments have set up proactive programmes to sample food and water for the occurrence of microbiological hazards of concern, which can be determined as a percentage of contaminated samples (the prevalence) and/or the number of microorganisms, e.g. CFU/gram of food. In addition, governmental agencies (inspection and control services, or assigned laboratories) carry out routine surveillance monitoring. Such data can be useful for hazard identification and also for exposure assessment or risk assessment. Most pathogen testing is presence/absence testing, because of the low expected contamination, and usually involves sample enrichment to allow the

target organism to grow enough to improve detection. Thus these tests are non-enumerative, unless multiple samples are tested in which case the proportion of samples that become positive can be used to estimate the concentration, similar to the Most Probable Number (MPN) method (e.g. Kiermeier *et al.*, 2011). There are some hazards for which tests do not yet exist, so even prevalence data may not be easily obtained. For example, until relatively recently, no reliable diagnostic tests were available for norovirus. This situation has now been addressed using molecular methods, though it still is not yet possible to differentiate between infective and non-infective virus particles (e.g. DNA fragments and damaged capsid). Finally, it should be noted that the efficacy of testing frequently depends on the size of the analytical unit tested, e.g. 1g versus 25g (Funk, Davies and Nichols, 2000; Vimont *et al.*, 2005).

In many of the exposure assessments published to date, the lack of specific data on primary production has often been identified as a weakness. Occasionally, governments or other stakeholders arrange programmes that specifically survey establishments involved in primary production. However, such programmes are often run for other purposes, e.g. to better understand pathogen ecology and production hygiene with the aim of improving or refining control measures when necessary. Such studies are often small, specific studies that typically concern one hazard and one commodity type (e.g. *Salmonella* in broiler chickens), which may nevertheless be enough for a specific risk assessment.

If national data on foodborne pathogens are not systematically collected in a country or region, it may be possible to utilise data from another country. In that case, the rationale for the choice of country and information on the possible limitations of the data in representing the current situation in the country in question need to be clearly documented.

Strengths

These activities generate substantial amounts of data, both in the form of prevalence or contamination level information. The potential for the use of such data in exposure assessments should be good, especially for systematic monitoring that covers a wide range of products in a certain category and a significant area (a country or region). To allow optimal evaluating of data on prevalence and level of contamination, proper descriptions of the details (i.e. year, season, geographical location, country, etc.) should be provided.

Limitations

Surveillance data collected by different government agencies are rarely pooled and the raw data may not be readily available or easy to obtain. Also, a detailed description of the product or hazard may not be provided. Additionally, a major drawback is that these data may not be random or fully representative. They are generated as part of official control systems that often take account of resource limitations by targeting foods that are known to be problematic. Alternatively, they are generated to support food inspection processes where samples are only taken if there appears to be something wrong with the hygiene of the premises or process, hence these data are often biased. In many cases, the lower (and upper) limit of detection (LoD) and analytical unit size are not reported, and neither are the sensitivity and selectivity (or specificity) of the detection method(s) utilized. Surveillance data collected at both primary production and processing/retail have a clear limit in terms of geography and time.

10.2.1 *National food production statistics*

Food production statistics provide an estimate of the amount of food commodities available to the total population, and as such can be useful for exposure assessment. Examples of this type of data

include the FAO Food Balance Sheets (FAOSTAT[15]) and other national statistics on total food production, disappearance or utilization. Because these data are available for most countries and are compiled and reported fairly consistently across countries, they can be useful in conducting exposure assessments at the international level.

Strengths

These reports contain detailed information and provide a good overview of a country's production of food commodities and imports.

Limitations

Figures reported may be outdated and for some food commodities, production statistics may not be available. It is important to note that production statistics are not necessarily specific to how much of the product is destined for the food supply as compared to other uses, e.g. biofuels. In addition, total amounts of a commodity may need to be adjusted to account for spoilage or other losses to arrive at the total amount that is consumed as food.

A 'reality check' relating food consumption to food production should be undertaken where possible. That is, if food consumption statistics are available and they are aggregated over the whole population, does the total amount of the food consumed approximately equal the total production for food, considering imports and exports, (likely) losses during processing and preparation and general wastage? If not, then some of the assumptions underlying the calculations may need to be critically assessed.

10.2.1 *National consumption databases*

Two types of food consumption data are frequently used for characterizing food consumption patterns for MRAs: food production statistics and food consumption surveys. These data can be very useful in exposure assessments. Other sources of information such as retail food sales or purchase data may be useful in filling data gaps in either food production or food consumption survey data. When using such data, allowance should also be made for the effects of food wastage and food spoilage.

Some countries have carried out Food Basket studies to describe the amounts and frequency of foods consumed. In countries where household food surveys have been carried out, useful information for exposure assessments might be available. In addition, the use of 'Participatory Epidemiology' methods (Mariner and Paskin, 2000) could be of value in data collection as well, being based on participatory techniques for gathering information based on community observations and traditional oral history (Bergold and Thomas, 2012).

Another data source of potential use is the WHO GEMS/Food consumption database[16]. This database provides information for a total of about 500 items at up to three levels of statistical food categorization on a country/cluster basis. These data may provide a useful starting point, though care needs to be taken with respect to interpreting the results. Where possible data should be checked against other sources.

Food consumption patterns will probably differ based on population demographics (age, gender, ethnicity, health status, socioeconomic group) and seasonal and regional (both national and international) differences in food availability. Consideration of food consumption patterns for sensitive subpopulations (e.g. young children, pregnant women, the elderly and the

[15] http://www.fao.org/faostat/en/#data accessed 12 July 2019
[16] https://www.who.int/nutrition/landscape_analysis/nlis_gem_food/en/ accessed 10 December 2018

immunocompromised) and high-risk consumer behaviour (e.g. consuming unpasteurized dairy products or undercooked or raw meat products) are particularly important. Information that enables estimation of variability in serving size will also be important.

Strength

Food consumption surveys can provide detailed information regarding the types and amounts of foods consumed by individuals or households and sometimes also the frequency with which the foods are consumed (van Rossum *et al.*, 2011). These surveys usually include a representative sample of individuals or households, from which consumption for the total population or specific population subgroups can be extrapolated. It is possible that food consumption data may be available for the 'at risk' group for a specific area.

When surveys are repeated over time then changes in consumption patterns may be observed.

Since serving size directly affects the numbers of pathogen consumed, these surveys may provide a method to determine a distribution of amounts consumed. Although the surveys are usually short in duration (one or two days to a week for each survey participant or household), they provide detailed information about the types of food consumed, as well as when and where foods are consumed (van Rossum *et al.*, 2011).

Limitations

Food consumption patterns may vary widely within a country and the consumption estimates derived from national food balance sheets will not reflect this variability. For example, in Sub-Saharan Africa the majority of the population live on the land and eat what they produce, though there may be considerable differences in consumption from the population that lives along the coastal areas. National food consumption surveys would be of great value here, but they are conducted in relatively few countries worldwide.

Not all national survey data sets have raw data by time of day and place of consumption as well as a total amount of each food consumed, and even if they do, it is often difficult to extract this type of information and analyse it (e.g. the time of day needs to be clearly defined at the time of the survey, as well as when data are sub-divided for analysis, etc.). It also requires fairly sophisticated software to be able to analyse individual dietary data at this level of detail, as opposed to deriving mean or median population statistics. This is particularly true if all sources of a food are required to be aggregated at an individual person level (e.g. apples from raw apples, apple juice and apple pies). In terms of microbiological risk assessment, this addition of food consumed from different sources also has additional problems as each food source is likely to have a different level of contamination of the hazard due to different food processing and preparation routes.

Food consumption surveys generally do not record descriptive information about the foods that may relate to food safety. For example, they may not report whether milk was raw or pasteurized, whether a soft cheese was made from raw milk, whether cooked shrimp were domestically produced or imported, or whether a food was packaged by the processor or at retail. For this information, food sales data from industry, trade associations, retail stores and other sources can be combined with results of food consumption surveys to estimate the frequency with which very specific food products might be consumed. Whenever possible these data should be compared with information from epidemiological studies (case control, cohort or outbreak investigations) to verify or calibrate that food survey data capture the actual risk factors.

10.2.2 *National population census*

Governments regularly publish reports on population size by region, gender, age strata, etc. These figures may be useful when characterizing the risk at population level and/or by type of population.

Strengths

These reports contain detailed information and provide a good overview of the country's population demographics, including age (possibly group, e.g. 20-25 years, etc,), gender, socio economic status, etc.

Limitations

Reports on population census may be outdated as censuses are very resource intensive and are therefore undertaken relatively infrequently, e.g. every seven or ten years. Care should be taken to ascertain how the census was administered and what sub-groups of the population may not have been captured, e.g. homeless people, and what approaches, if any, have been used to adjust for these.

Population statistics generally do not capture at risk groups, unless they are specifically related to demographic characteristics, e.g. age.

A specific problem for international exposure assessments is that information and data may not be accessible due to language barriers. Both finding relevant data and correctly interpreting their context may be a problem.

10.3 Industry data

Both textual and numerical data can be obtained from industry stakeholders, including occurrence of microbiological hazards, production stages and processing conditions, description of the final product and product pathways. Data on product sales and market share may also be available from private marketing agencies, trade associations and industry. These data are very relevant for the exposure assessment.

Industry can furnish information on whether the product is fresh or frozen, whether it is sold cooked or uncooked, whether or not it is further processed and the extent to which ingredients are mixed. A complete description of the food, including salt levels, pH, packaging and other relevant information should be provided. Such data may also refer to other factors that may influence the prevalence and/or concentration of hazard in the food, e.g. the extent to which the product and sub-products are domestically produced or imported; the different ingredients added; or other products typically consumed with the product.

The food chain consists of all stages from primary production to the consumption (including home, restaurant, foodservice, and/or institutional locations), and thus data relating to each of these stages are required as part of the exposure assessment. Using meat processing and distribution as an example, the various stages will include the farm; transport to and holding at a slaughterhouse or processing plant; slaughter; processing; packaging; storage; distribution and retail; transport to the home; handling and home storage; food preparation; and consumption. Some of these stages and processes may vary between producers, retailers and consumers and thus it is important to obtain information to describe and account for this variation – this is particularly pertinent for exposure assessments where formal and informal supply chains exist. Certain stages or processes may be regulated, for example, with respect to the use of chemicals or additives; such regulation and information on the extent to which they are followed in practice may give relevant data to be collected.

Considering growth and survival of a microbial hazard, the times, temperatures, and other ecological factors such as pH at the various stages are important. Particular examples of requirements include the duration of, and temperature during, storage or transport; freezing temperatures; pasteurization times and temperatures; cooking times and temperatures; and the addition of ingredients that may alter pH. Data that enable description of the variation in these parameters, for example from producer to producer or day to day, are also important. Often, individual stages in the food chain are considered to be static for a specified period. However, certain conditions, such as temperature, are more likely to be cyclic and data should reflect that. While data may be readily available on thermal inactivation, data on other types of thermal or non-thermal processing that affect microbial growth and survival may not be as readily available.

It is also important to gather information relating to the stages of mixing and partitioning. For example, the meat from an individual beef carcass can be partitioned and then perhaps mixed with meat from other beef carcasses to produce a ground beef burger. Partitioning and mixing will influence the microbial status of the product, in terms of both likelihood of contamination and number of organisms, and thus data that are descriptive of these processes should be collected. Typical requirements will include the extent to which both events occur, the numbers of carcasses or products contributing to a mixed product, and characteristics of products obtained through partitioning (including distributions in quantity and size).

Retail surveys also represent another source of industry data, including information on geographical area, season, and the degree to which the data represent all manufacturers, distributors or retailers.

Strengths

Industry collects vast amounts of product/commodity specific data, which it stores in a wide array of private systems. Gaining access to such data and information about product pathways, can provide the risk assessor, and risk manager, with important information about the realities of the food production, that might otherwise not beknow with confidence.

Where sampling, testing and monitoring programs are in place, information will be available over time and at various stages during the production, from supplied raw ingredient through to finished product. Such data will be useful for application of predictive microbiological models when the fate of microorganism is to be predicted.

In some businesses sampling and testing of the raw material and end product is extensive and frequent as it is the primary means of 'ensuring' food safety. Other businesses employ a preventative approach to food safety such as the implementation of a food safety management system based on the principles of HACCP. In these businesses microbiological testing is infrequent and solely for the purpose of verifying the effective working of the HACCP system. Furthermore, the food production environment is sampled due to considerations of cross-contamination

Limitations

Major limitations to the inclusion of industry data in exposure assessments are the facts that they may not be hazard specific and are difficult to combine when generated in industrial settings that are difficult to compare individually. Because sampling and testing is usually done for verification purposes or to satisfy regulatory requirements, the data often concern the presence/absence of a microbiological hazards rather than the levels/concentration. When testing is done for indicator organisms the levels of contamination are usually recorded, but mainly for generic groups of microorganisms, such as total viable counts or Enterobacteriaceae.

In addition, access to and mining (retrieval) of such data is a problem in practice. In this regard, there is also a need to address the issue of confidentiality, which may be a stumbling block in relation to access. The use of proprietary information and data can poses some challenges and how to keep this information and data confidential needs to be discussed and agreed prior to their provision, as is done by FAO/WHO (2018b).

In addition, potential biases need to be considered, especially relating to the difference in processing and food safety programs related to business size. Large manufacturers process on an industrial scale with better and more automated equipment than small food producers. As a result, they can supply more geographically diverse retailers and supermarkets and access a different consumer segment than small food producers who sell their products at informal markets (especially in developing countries).

Similarly, large enterprises are more likely to have food safety programs in place, including spending (more) money on microbiological testing programs. This contrasts with small or very small enterprises, which are less likely to undertake much, or any, microbiological testing. If a food is tested for microbial contamination, anywhere in the food supply chain, then the industry stakeholders should provide sufficient information on the food, microbiological methods, sampling design and frequency of sampling, etc. However, such information may not readily be available.

An important limitation in collecting retail survey data by any group (e.g. trade association, academia, consultant) is that the identification of a contaminated food might trigger a recall (e.g. *Listeria* in a ready-to-eat (RTE) food or *E. coli* O157 in ground beef). This may make such surveys of limited value because any kind of recall may change the foods in distribution and impede future industry cooperation. Alternatively, when such studies are commissioned by industry, there may be a limitation placed on the type of microbial data that is collected, e.g. hygiene indicators and/or presence absence of specific genes rather that direct isolation of the pathogen.

10.3.1 *Description of product and supply chain*

Throughout the food chain, many control options are available to reduce the risk of microbiological contamination of the final food product. These may be incorporated in HACCP plans that are specific for each product and manufacturing site, and thus may vary substantially between manufacturers. Data should be collected that describe both the methods of control and the extent to which these vary. Examples include cleaning and disinfection methods and the extent and frequency with which these are undertaken; inactivation methods and their critical limits; any testing of live animals and intermediate or final products, with estimates for test sensitivity and specificity; and handling practices.

10.4 Unpublished data

Potentially vast amounts of data (generated throughout the world) are never published in a form that can be used by others. This can be due to many different reasons that can for instance relate to the attractiveness of the subject to publishers or the (scientific) community, i.e. publication bias, barriers in communication (resources, language) or due to time and/or resource constraints for the researcher. This is an unfortunate situation; such data could give new insights, reduce uncertainty and avoid unnecessary duplicate experimentation. However, like other data sources, the quality of unpublished data need to be ascertained carefully before use in a risk assessment.

Some steps can be taken towards improving access to such data. Building networks is very important in this regard, as these can be used to inform a wider audience of the data needs for risk assessment and also provide a means of gaining information about, and even access to, unpublished studies.

Building up a relationship with potential data providers is essential in establishing trust and instilling confidence that the data will be used properly and remain confidential if necessary. There is a need for networking, especially with others who might be working in areas where data are required.

Another avenue for gaining access to unpublished data is through public calls for data, which is usually undertaken for international risk assessments by FAO/WHO, as well as national competent authorities. This also forms an important part of risk communication and helps to involve different stakeholders.

10.5 Data gaps

All risk assessments require data and knowledge (of processes, interactions, etc.), irrespective of whether they are qualitative or quantitative. Data and knowledge gaps influence the assessor's confidence in the risk characterization and the robustness of the estimate. The form of a risk assessment is determined primarily by looking at what decision questions need to be answered, taking into account the decision criteria described in Section 3.5. Then a search is done to see what data and knowledge are available that would help construct a logical risk-based argument (the risk assessment) that answers these questions. A balance is generally needed: taking a particular risk assessment approach may not be able to answer all questions but may provide a better-quality answer. Data may not be available to answer the question at all. Thus, defining the form of a risk assessment may require considerable dialogue between assessor and manager.

Both numerical and textual data are required to model all stages of the exposure pathway. Often, data are limited or non-existent. However, a lack of knowledge about a process should not necessarily inhibit the ability to conduct an exposure assessment. When deficiencies in the data exist, they must be clearly communicated to the risk managers and documented in the exposure assessment. Such communication will ensure that additional data requirements are identified. Even in situations where appropriate and representative data are known to exist, problems can still occur. For example, there may be institution or company confidentiality to consider, the data may be politically sensitive or there may be a charge for use of the data. The iterative nature of risk assessment allows for the continuous upgrading of data as new information becomes available.

This process will often lead to a better understanding of the value of other information that is not currently available. One can ask what else could be done if some specific data could be found. Depending on the time left until a decision has to be made, and on the resources available, the risk manager may consider it worth waiting, or expending the resources to acquire those data and hopefully be able to make a more informed judgement as a result.

It is tempting to plan out the structure of a risk assessment that will answer all the risk managers' questions, and then attempt to find the data required to 'populate' the risk assessment. However, in the food safety area this may not be a practical approach. Food safety management is beset by a lack of data, so writing a wish list of all the data one would like will inevitably lead to disappointment. Other approaches, such as building simplified model-based reasoning to describe the system or process before considering the data availability (Ebel *et al.*, 2012), have been proposed as preliminary activities to aid in determining the form of the risk assessment. A brief list of reasons for such gaps includes:

- it has not previously been seen to be important to collect these data;
- data are too expensive to obtain;
- data are impossible to obtain given current technology;
- past data are no longer relevant;

- data from other regions are not considered relevant; or
- the data have been collected or reported, or both, in a fashion that does not match the risk assessment needs.

Data that have not previously been seen to be important often arises in contamination studies with infrequent detection data. Such data are not usually valuable for scientific journals; therefore, researchers have less interest in conducting such studies. However, data on non-detections are important for risk assessment, e.g. to estimate prevalence.

Using the risk assessment framework, it may be possible to determine which gaps have the most influence on being able to address the risk management questions. This identification process can be used to set priorities for future data collection and experimental research.

There are a number of approaches that can be used to help overcome limitations in data. These include model design, surrogate data, expert opinion and the collection of new data.

10.5.1 *Model restructuring*

Ideally, all stages in the exposure pathway that affect the hazard are included in the model structure. However, in many situations, data for specific stages may be limited or even non-existent. Also, the statement of purpose for conducting the risk assessment may not require detailed analysis of all processing stages, i.e. a farm-to-fork exposure assessment may not always be required. When this is the case, it may be possible to restructure the model to exclude the stage for which data are not available or in such a way that alternative available data can be used (e.g. beginning the exposure assessment after the processing stage and obtaining prevalence and concentration using monitoring data); clearly any changes in the scope must be discussed and agreed with the risk managers. In addition, simplification of the model may have the benefit of reducing the compounding of uncertainties. There are limitations with this technique, as important factors that have an effect on the risk may be overlooked and lead to errors. Cullen and Frey (1999) provide a useful discussion of trade-offs regarding various levels of model complexity.

10.5.2 *Surrogate data*

In one sense, nearly all data are surrogate data unless specifically collected as part of the exposure assessment. Pilot plant data, for example, is a surrogate for production facilities; thermal death time values obtained via capillary tubes are surrogates for inactivation in the plate pasteurizers used in food processing. Classically, certain benign species or strains of microorganisms are used as surrogates for pathogenic strains. In such cases, the relevant characteristics of the surrogate organisms should be the same as the organism of interest, or the differences documented and taken into account. Surrogate organisms are more appropriate for quantifying or predicting treatment efficacy than for predicting or quantifying health effects such as actual dose-response relationships. The appropriateness of the surrogate data must be judged when assigning uncertainty to the data. For transparency, use of surrogate data must be described and justified.

Indicator microorganisms for particular microbiological hazards have been used in some exposure assessments where data on the hazard is not available or cannot be collected. An example would be the cross-contamination rate of *E. coli* O157:H7 from faeces to animal carcasses. Because of the low prevalence of *E. coli* O157:H7 in faeces, a direct measure of contamination cannot readily be obtained. The easily measured generic *E. coli* is therefore used as an indicator of faecal to carcase transfer, which can then be related back to *E. coli* O157:H7. When using surrogate data, care should be taken to clearly identify where it was used and any underlying assumptions (e.g. proportionality between the pathogen and surrogate) should be explicit whenever possible.

Regarding food consumption data, if there is insufficient detail to provide estimates for at-risk populations (pregnant women, immunocompromised, elderly, etc.), data for comparable age and gender groups in the normal population may be used. Data from other countries or regional data may also be used for food consumption if it is known that food consumption patterns are similar.

Sensitivity analysis (Chapter 15) of the final model can be used to determine if the parameter, for which surrogate data were used, has a significant effect on the final risk. If the parameter is important in estimating the risk, then an additional study may be undertaken to try to collect more relevant data.

10.5.3 *Expert knowledge elicitation (EKE)*

Expert knowledge elicitation is a formal approach to the acquisition and use of expert opinions, in the absence of or to augment available data. It will inevitably be necessary to elicit expert estimates for parameter values in the model where there is a critical lack of data, and where for pragmatic reasons it is essential to assess that risk in the relatively near future. Problems here include, for example, decisions on identification and selection of experts, the number of experts required, techniques for eliciting information, overcoming bias, etc., and methods are still being developed in this area (e.g. Jenkinson, 2005; Hemming *et al.*, 2018; Dias, Morton and Quigley, 2018).

Such expert opinion should be elicited using formalized and documented methods that avoid bias and can be used to formulate appropriate probability distributions (Gallagher *et al.*, 2002; Nauta *et al.*, 2001; Vose, 2008). In situations where expert opinion differs markedly, weighting methods can be used to integrate information in the most reliable manner. Experts should strive to transparently document the rationale supporting their opinion to the greatest extent possible.

When expert opinion is required, the problems and methods of selection, overcoming bias, etc., up to this point are likely to be similar irrespective of the level of quantification used for the risk assessment. It is accepted that ideally a 'sufficient number' of experts should be utilized. Techniques like the Delphi method (Linstone and Turoff, 2002), and modifications such as IDEA (Burgman, 2015; Hemming *et al.*, 2018), which aim to achieve consensus among a panel of experts, can help produce more reliable estimates from the available information. However, there are situations when there truly are very few, and on occasions perhaps only one, expert in the specific topic worldwide. Sometimes there are no true experts. This leads to the use of inputs with very wide levels of uncertainty, whatever the risk assessment type, which is far from ideal but may be the only option in the short term.

In a quantitative risk assessment, it is necessary to convert expert opinion into a numerical input, and once again various methods exist and are being actively developed (e.g. Gallagher *et al.*, 2002; Burgman, 2015; Dias, Morton and Quigley, 2018). Even in a qualitative risk assessment, these methods may be used to convert expert opinion into numerical values for specific model steps and this is, where time allows, the preferred method. As noted earlier, when used to describe approaches to risk assessment, the terms quantitative or qualitative do not refer to formally defined categories of risk assessment. An alternative and less sophisticated way of using expert opinion in qualitative risk assessments, however, may be to ask directly for an opinion on the probability of a specific step in narrative terms of, for example, high, low, negligible, etc. The meanings of these words will have the same subjectivity problems as those discussed for qualitative risk assessments in general (see Section 7.2), and the reader's evaluation of the results will need to be based on their evaluation of the experts selected. In principle, such a method should be only a temporary measure until improved data are available.

The estimation of dose-response model parameters is unlikely to be based on expert elicitation and instead based on model fitting. The choice of dose-response function, that is, the mathematical form, is often based on the modellers' expertise, and thus forms a type of expert opinion. When no dose-response model exists the likely dose needed to result in a specific human health effect, e.g. ID_{50}, may require expert elicitation; especially for emerging hazards that have not been studied extensively.

Readers with further interest in the use of expert opinion should consult Morgan and Henrion (1992), who present a sequence of chapters summarizing the heuristic biases in expert elicitation, a typical formal expert elicitation protocol intended to overcome such biases, and examples. Informal EKE can be performed with less experts without the presence of an experienced facilitator in the sense that a small group of scientists wish to quantify their own knowledge about an uncertain quantity, for the purposes of some scientific endeavour. In any case however, the scientists' judgements should be made as carefully and objectively as reasonably as possible and documented fully, according to the principles of the formal EKE. Additionally, EFSA Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment (EFSA, 2014a) and the Intergovernmental Panel on Climate Change (IPCC, 2001) discuss the process of expert elicitation in scientific assessments.

Strengths

When there is a lack of the specific data needed, say, to develop dose-response relations, but there are scientific experts with knowledge and experience pertinent to the elucidation of the information required, expert elicitation provides a means of acquiring and using this information. This can involve the development of a distribution for a parameter in a model for which there is no, little or inconsistent numerical data, through the use of accepted processes that outline the lines of evidence or weight of evidence for generation of the opinion and use of the results. It is generally not expensive, particularly in relation to short-term needs.

Limitations

The results obtained depend on the methodology used and are inherently subjective and thus open to debate. The results also depend on the experts selected and may have limited applicability for issues involving an emerging hazard.

10.5.4 *Collection of new data*

At times the need arises for the collection of new data, e.g. prevalence and concentration data for a foodborne pathogen at a specific point of the food chain. The process of obtaining an estimate of the prevalence or the enumeration of microbiological hazards usually involves the following steps:

1. Define the research question
2. Identify the reference population and study population and obtain an appropriate sampling frame
3. Design a sampling scheme and identify the sample population
4. Collect and analyse appropriate samples
5. Conduct statistical analysis of the data

Those contemplating the collection of new information for use in risk assessments should consult a statistician or someone trained and experienced in data collection, especially someone who is familiar with the underlying research domain, e.g. microbiology, consumer behaviour, etc.

10.6 Recommendations on data collection and organization

The characteristics of the data that might be needed at a particular stage are likely to vary from assessment to assessment. Whilst certain characteristics may be considered ideal, in practice it is often necessary to use, in the first instance, whatever data are available. This brings into focus the iterative nature of a risk assessment, which is concerned with the fact that initial attempts to model a process are likely to utilize data with a high degree of uncertainty. This process can be used to identify where the greatest uncertainty lies, allowing targeted data collection for subsequent model updating. Gradually, with further iterations of the modelling process, the uncertainty is reduced. Thus, the first iteration of the assessment might be undertaken specifically to identify data needs and/or data gaps. The second iteration may assess the risk of exposure, but with wider uncertainty limits; and the third iteration, using 'new' data, may allow an estimate of the exposure with a narrower uncertainty band and higher predictive ability. There may be considerable time delays between these stages. The level of uncertainty should be included in the data description.

10.6.1 *Searching for data*

Search protocols using computer-searchable literature databases and data repositories, such as,

- Promed: http://www.promedmail.org/
- Pubmed: https://www.ncbi.nlm.nih.gov/pubmed/
- ComBase: http://www.combase.cc
- FAOSTAT: http://www.fao.org/faostat/en/
- WHO/GEMS: https://www.who.int/nutrition/landscape_analysis/nlis_gem_food/en/
- FoodRisk.org: http://foodrisk.org/
- Food Science and Technology Abstracts (FSTA): https://www.ifis.org/fsta
- OVID Current Contents: http://www.ovid.com/site/catalog/databases/862.jsp
- Web of Science: https://clarivate.com/products/web-of-science/
- Scopus: https://www.scopus.com/

should be devised that are comprehensive and reproducible but are also appropriately selective. Systematic plans for obtaining literature that predates these databases or that is not indexed in them need to be devised using citations in more recent publications, reviews and book chapters. Criteria for search protocols and data selection should be transparent, with appropriate explanation recorded in the documentation.

The Research4Life resources noted in Section 3.5.2 may be helpful in gaining access to scientifically published articles and associated data.

10.6.2 *Selection of data*

It is frequently stated that 'all data are biased'. Nevertheless, data should be as representative as possible of the food, microbial or process parameters being assessed and the population consuming the food. Preferred data generally comes from peer-reviewed publications, followed in importance by non-reviewed or unpublished data (government documents, theses, proceedings, etc.; see Chapter 10 for details). Some data are not available in the peer-reviewed literature (e.g. consumption data), and it should be remembered that even peer-reviewed data are, in most instances, not collected for the purpose of being used in exposure assessments, and thus may not comply fully with all data requirements or be fully representative for the case at hand. Any biases or limitations in the degree to which data represent any particular point of view should be identified and documented (e.g. funding source). When no or too few data are found, expert opinion will need to be used (see Section 10.5.3). Generally, the data should be as close as possible to, or specific to, the requirements of the exposure assessment. For example, if the exposure assessment were to

calculate the exposure in a particular country, the preferred data would come from that country. The next choice would be data in that region or a comparable country; the final choice would be from somewhere else in the world (keeping in mind the purpose of the risk assessment). Selection criteria should include consideration of factors such as geography, time, microbial strain, methodology, equipment type and design, and population demographics. Food consumption data should provide sufficient detail to allow estimates of consumption of the food(s) of interest per meal or per day. The data should be representative of the total population, and ideally will provide information about subgroups within the population.

10.6.3 *Formatting of data*

The ideal format of the data will vary with the particular type of data required; there is no one ideal format for all data. In particular, data that are descriptive of the biological and manufacturing processes will generally be textual, whereas parameter and model input data would, wherever possible, be numerical.

However, there are some underlying principles that should be considered when formatting data:

- Data should be fully referenced as to source (within the confines of commercial sensitivity).
- Units should be given where appropriate.
- Raw data, rather than average or other summary statistics, should be used wherever possible.
- When raw data are not available, a description of the distribution, the level of uncertainty and the amount of variability should be included to the greatest extent possible.

10.6.4 *Level of detail recorded*

When collecting data for use in an exposure assessment, it is useful to record and report detail to the most complete level available. This should be done in an appropriate way which does not interfere with the flow of the report to the extent that it hampers clear communication. The additional information that describes the data set is often referred to as *metadata*, and there are a number of metadata standards available, though not specifically for microbiological data. Examples of some details that might best be routinely recorded and reported are:

- Information on data source or provenance. This should include: the full reference for the source if a paper or similar; the name of the provider if a personal communication or unpublished data; the date of the collection of the data; affiliation or funding source, or both, of the data provider.
- Information on the study itself. This should indicate whether it was a laboratory- or field-based study.
- Details of sample, including: livestock species (giving scientific name where appropriate) or product definition; source (country, region, category of producer, chain of retailer, etc.); selection method (in particular for livestock, whether samples are clinical cases or random selection); population size; season of collection, if appropriate; portion description or size, if appropriate; and method of collection of samples.
- Information on microbiological methods. This should include: sampling method, microbial species, subspecies, strain, in as much detail as is available (and for pre-specified exposure assessment, the required detail should be specified and collected); tests used, including any variation from published methods; test performance characteristics; units used; and precision of measurement.
- Information on the results obtained. This should be recorded as the raw data, and include: number tested, together with results (including units) given for all samples tested.

10.6.5 *Combining data from different sources*

Representative data are oftentimes limited, and it is often preferable to use all of it. However, decisions need to be made when different data sets have different degrees of applicability and relevance to the parameter being modelled. Techniques such as meta-analysis (Petitti, 2000) can be used for the purpose of combining data sets. Gonzales-Barron et al. (2016) provide an overview on how to integrate prevalence data of pathogens from different sources. More generally, Bayesian approaches (Gelman *et al.*, 2013; Kruschke, 2014) may be useful when considering existing knowledge in the light of recently collected information. In certain situations, using Bayesian techniques allows a better estimate of the parameter to be obtained than if the recent data were used in isolation. When a data set is biased, the data may be adjusted before being combined with other data or used in the risk assessment. An example would be when recent research or methods development demonstrated that data collected by one method consistently underestimated the true parameter value by a known amount.

Weighting is often employed so that data sets considered more relevant have more influence on the estimated parameter value. Weighting by the number of samples is frequently used, so that larger studies have more influence. Weights may also be used to reflect the expert's belief in the quality and appropriateness of the data. Older data or data from another geographical area might be used in estimating the parameter value but be given less weight. The selection of the numerical weighting factors is highly subjective and should be explained for full transparency. Composite data sets may be obtained by averaging, method of moments (Hansen, 1982), or maximum likelihood estimates. Careful examination of the different data sets may facilitate estimates of variation (e.g. different microbial strains used in different studies) or uncertainty (residual errors in statistical analyses). Meta-analysis and mixed-effect models can also be used to evaluate data variation.

To avoid inserting the risk assessor's biases into the parameter values, data should not generally be ignored or deleted. However, certain data sets may clearly be inconsistent with the greater collection of data and knowledge. Comparing the size of the remaining distribution with the divergence of the particular data set may suggest that a particular data set should be excluded. This should be done with caution, as the outlier may indicate another source of variation that is otherwise being overlooked (see Figure 18 in Section 16.1).

10.6.6 *Presentation of data*

The format of the data will affect the method of presentation. The underlying principle is that the presentation should be clear and easy to follow. Again, the data may be textual or numerical. When presenting a large amount of data for a particular exposure assessment, a contents table or list is desirable. An introduction or overview of the assessment puts the data to be presented in context. The data should then be presented in a logical order.

In general, with an exposure assessment, there are one or more pathways by which the consumer may be exposed to the microbiological hazard. The first part of the data to be presented is generally the textual data that describes these pathways. For complex pathways, a high-level overview of the process may be required, followed by a more detailed description for each step in the pathway. Also, graphical presentation of the pathways, such as in the form of a flow chart, is generally helpful.

When presenting numerical data, this should also follow a logical order, and this is again likely to follow the order of the steps in a particular pathway. A tabular format is frequently useful, particularly for raw data. However, enough text should be provided to fully describe the relevance of the data, and how they are utilized in the assessment. Summary data are often also best tabulated. Graphs or histograms may in addition be used to clarify data but should not be used without

explanation. Titles of tables or graphs should allow them to be fully identified and should be unambiguous. References should be clear within the text, diagram or table, and a comprehensive reference list given. Any web pages or similar are probably best attached as appendixes.

11. Quantitative modelling approaches

As introduced in Section 5.2.3 there are different categories of quantitative models. The following categories increase in model complexity and thus also increase in the potential richness of the model outputs. This is achieved by incorporating variability and uncertainty into the model and this allows their effects, and those of the model inputs, on the exposure or risk outputs to be evaluated (see also Chapters 14 and 15).

11.1 Deterministic

Deterministic models assume that inputs to a model are known and fixed values with no variability or uncertainty. Although they are simple models they generally require more data than for a qualitative assessment. A single value, e.g. average, highest level, most often observed value, 95th percentile, etc., is chosen to characterize each input variable in the model such as the concentration in the food; the log reduction from cooking; the amount of food consumed per serving, or the frequency of consumption; etc. The individual point estimates are combined using mathematical models to generate a point estimate of exposure, and, through a dose-response model, the consequent risk. An example of a deterministic model, implemented in a generic framework, is RiskRanger (Ross and Sumner, 2002; Sumner and Ross, 2002). The effects of changes to model variables can then be investigated by 'what-if' testing to generate outputs. For example, the initial scenario may be based on the average for each input variable. Subsequently, however, the difference in the risk estimated from using the most likely value compared to the 95th percentile value, and other scenarios, could be investigated.

When conducting deterministic exposure assessments, selecting a conservative value for each variable has often been used to develop deliberately conservative, 'safe' or 'worst case' estimates. Propagating such conservatism through the model, however, can result in an unrealistic over-estimate of exposure because the exposure estimate can be based on a highly improbable scenario. Thus, a drawback of the deterministic approach is that the likelihood or probability of the estimated exposure actually occurring is unknown. Some values are more likely to occur than others, and without knowledge of the likelihood of each outcome, the risk manager may inappropriately allocate valuable resources to reduce an event that rarely occurs. Stochastic models can overcome this problem.

11.2 Stochastic

The stochastic, or probabilistic, assessment represents all the information available for each input variable, which is described as a probability distribution of possible values. Most parameters such as pathogen prevalence in primary production, pathogen concentration and growth, storage temperatures, serving size, and storage times have a range or distribution of values. These variables are better described as distributions so a realistic range and frequency of values can be represented. In stochastic models, scientific data are used to generate and define probability distributions for each input variable. They are then combined to determine the probability distribution of an adverse outcome (Ruzante *et al.*, 2013). Consequently, the outcome of a stochastic exposure assessment is a statistical distribution that describes both the range of doses of the hazard that might be experienced by an individual or population, and the likelihood of each level of exposure. For example, consider a hazard in a food product.

- The concentration of the hazard in the food prior to heating is log-normally distributed with mean and standard deviation of 1.0 and 0.8 $\log_{10}$ cfu/g and
- The effective reduction from heating the food is also log-normally distributed with mean and standard deviation of 2.5 and 0.7 $\log_{10}$ cfu/g.

Analytically it can be determined that the concentration of the food after heating is also log-normally distributed, with mean 1+(-2.5) = -1.5 $\log_{10}$ cfu/g and standard deviation of $\sqrt{(0.8^2+0.7^2)}$ = 1.06 $\log_{10}$ cfu/g.

However, finding analytical solutions, as above, for a stochastic exposure assessment, often involving numerous stochastic inputs, is usually not possible, particularly if the distributions are not 'Normal'. For this reason, Monte Carlo simulation is usually used to perform the assessment (see below).

The distribution used to describe a data set is dependent on the number and pattern of data points available, and on the knowledge about the nature of the phenomenon or process being modelled. Detailed reviews of the important probability distributions are available in the literature (Cullen and Frey, 1999; Haas, Rose and Gerba, 2014; Morgan, Henrion and Small, 1992; Vose, 2008). Uncertainty in parameter values can also be expressed by probability distributions, as discussed in Chapter 12.

The transition from qualitative assessment to deterministic assessment to stochastic assessment usually represents an increase in both information and time required. However, due to the availability of simulation modelling software, the time involved for a stochastic assessment may not be much greater than for a deterministic analysis. Despite its increased computational complexity over the deterministic approach, much of that complexity is dealt with by the software and the stochastic method is favoured among most risk assessors because it generates more information to support decisions, e.g. by identifying the range of possible exposure levels from all possible exposure routes from which the most likely level of exposure, or any specified percentile value, can be determined. This output provides much greater information than a single point estimate. In addition, stochastic modelling allows for explicit identification, modelling and separation of variability and uncertainty (Chapter 12). However, with the increased complexity also comes the increased risk of introducing errors into the assessment, and the output is more difficult to understand, interpret and use for decisions.

11.3 Monte Carlo simulation

As noted above, stochastic models are generally complex in nature, and as a result are usually difficult, or impossible, to solve analytically. To overcome this problem, the model can be evaluated on a computer, using Monte Carlo simulation. A variety of specialized computer software packages are available to support this approach and are discussed in various texts (e.g. Cullen and Frey, 1999); a good summary is provided in Table 1 of Basset *et al.* (2012). Commonly used programs are spreadsheet add-ons, such as @RISK© and Crystal Ball©. Microbial risk assessors have also used the stand-alone package called Analytica® or the US Food and Drug Administration's web-based, and free to use, FDA-iRISK system (https://irisk.foodrisk.org). Other mathematical (e.g. Matlab) or statistical packages (e.g. SAS, R) can also be used for simulation modelling, including various free add-ons, such as the mc2d package for R (Pouillot and Delignette-Muller, 2010). Models can also be constructed using general-purpose programming languages, including FORTRAN, Python, Visual BASIC or C. Commercial software packages may be less 'flexible' to use compared to programs developed in programming languages by the modeller, although both require specialist expertise to model the processes appropriately. Exchange of models may be hampered if the chosen software is not widely available, and open-source software packages that can be downloaded for free, may help to improve the ability for the risk assessment to be 'audited' by others. Simulation models that can be placed and run on the internet may also be desirable to further facilitate model evaluation (e.g. FDA-iRISK).

To undertake a Monte Carlo simulation, a mathematical model is constructed to describe the exposure assessment, including all variables that influence the exposure and their probability distributions. Collectively, the result of the combined equations is an expression of consumer exposure. The software then evaluates the model by generating, at random, a value for each variable from its corresponding probability distribution. The generated values are then combined according to the mathematical equations that comprise the exposure assessment model, and the exposure is calculated. A single realisation of this generation and calculation process is called an *iteration* of the model and represents the exposure from one possible combination of circumstances. There are many such sets of circumstances, however, some more or less likely than others and leading to greater or lesser exposure. To estimate the full range of possible exposures and the likelihood of each, the simulation software repeats the calculations many times: hundreds of thousands or millions of iterations are commonly performed. The result of each iteration is recorded and the distribution of exposures and probability of each is generated and forms the exposure assessment; intermediate results may also be recorded to provide insights into the model.

11.4 Other model classification schemes

In addition to the classification of models used in quantitative exposure assessment as deterministic or stochastic, other non-mutually exclusive classification 'schemes' might be encountered, *i.e.* the use of one description does not necessarily preclude an additional description from another classification scheme. Several common schemes are mentioned below.

Models can also be categorized as empirical or mechanistic. Empirical models simply describe data or relationships in a convenient mathematical form, without necessarily having an understanding of the underlying biological mechanisms. For example, a smoothing spline (de Boor, 2001) may describe a set of data points adequately, even though there is no biological basis for it. Mechanistic models have theoretical bases formed from the understanding of the behaviour of a system's components, e.g. binary fission bacterial growth. If correctly formulated, then, a mechanistic model should provide a good fit to experimental data and thus allow the interpretation of the response in terms of known phenomena and processes. In practice, exposure-assessment models will probably contain both mechanistic and empirical elements.

Estimates of exposure can also be viewed from a temporal perspective: they can be defined as static or dynamic. Static estimates relate to a particular point in time, e.g. the probability and level of exposure associated with a random serving of the food product, or the number of contaminated servings consumed per year. In contrast, a dynamic approach would consider the way in which exposure changes over time, for example, reflecting seasonality of exposure (Anderson and May, 1992; Bailey, 1975) or the increasing contamination of a processing line as time from last clean-up increases (Nauta, Van der Fels-Klerx and Havelaar, 2005; Zwietering and Hasting, 1997a, 1997b).

12. Predictive Microbiology

Predictive microbiology can play an important role in exposure assessment and is used to fill in data gaps that would otherwise require more extensive data collection programmes. Predictive microbiology, in conjunction with mathematical models describing various environmental factors, e.g. including storage time and temperature, pH, water activity, etc., can be used to estimate the final level/concentration of pathogens or spoilage organisms in the food. For example, while the number of pathogenic bacteria in food at retail may be available, the number in the food immediately prior to consumption is not. It may, however, be possible to model the number of pathogenic bacteria in the food immediately prior to consumption, considering the storage, preparation and cooking conditions.

Predictive microbiology also has limitations. Not all hazards that are of interest have been characterized – and therefore not all microbial kinetic parameters are available, uncertainties surrounding predictions are not always given, and predicted values may not truly represent the real world if models have not been validated. In spite of the limitations, predictive models remain valuable tools for exposure assessment of pathogenic microorganisms in foods. Detailed descriptions of the application of predictive microbiology in MRA can be found in Ross and McMeekin (2003) and Ross (2008).

12.1 Modelling microbial growth and inactivation

12.1.1 *Microbial ecology of foods*

The possible responses of most microorganisms in foods include stasis, growth or death. In general, viruses and protozoa ('parasites') are inert in foods, requiring a living 'host' to be able to reproduce. While they cannot grow, they can be inactivated by various treatments and processing steps. Similarly, prions are not infectious organisms but are proteins. While they also cannot grow in foods they may be inactivated by some treatments, although they are very resistant to denaturation.

Populations of microorganisms in foods may display stasis, growth or death, depending on the formulation of the food ('intrinsic' factors) and the processing, distribution or storage conditions ('extrinsic' factors). They may even display different responses at different times in a single unit of food because conditions can change during processing, transport, storage and preparation.

While each organism may have a qualitatively similar response to changes in temperature, pH, preservatives, etc., the magnitude and type of response (e.g. growth, death, stasis) to different levels of these factors is specific to the hazard in question. While pH, water activity and temperature are the most frequently cited properties and typically have the greatest effect on microbial behaviour, many foods will have additional properties with important consequences. These include the levels of fat, oxygen, phosphates, certain spices, organic acid anions (especially acetate, lactate, sorbate and benzoate), nitrite, ionic and non-ionic humectants (sugars, salts, etc.), and antimicrobials such as benzoate or sorbates. Food structure has also been shown to play an important role in influencing microbial behaviour in some foods (e.g. Wilson *et al.*, 2002).

To estimate exposure at the time of consumption, it will be necessary to model the cumulative effect over time of the food's composition (which may change over time) and processing or storage conditions on the microbiological hazard. In some cases, changes in microbial numbers during processing may occur as a result of cross-contamination, rather than growth or inactivation. Note that the same considerations may apply to microorganisms in water, whether recreational or for drinking and food preparation or irrigation.

4804 It is important to understand under what circumstances growth, inactivation or cross-contamination
4805 may need to be considered. In Table 39 are provided indicative values, based on expert opinion, for
4806 the effect of temperature on the rates of growth or inactivation of many vegetative bacteria;
4807 inactivation of endospores requires considerably longer time and/or higher temperature. Growth
4808 rates for fungi will be slower, but inactivation rates are generally in the same range.

4809 4810 *Table 39: Indicative response times for growth and inactivation of vegetative bacterial cells as a function of temperature.*

Temperature (°C)	Time for 10-fold increase in numbers (hours)	Time for 10-fold *decrease* in numbers (for vegetative cells)
-80		years to decades
-20		months
0	15-75	
5	10-30	
10	5-20	
20	3-10	
30	2-3	
35	1-2	
50	growth not possible for most	days to weeks
60		hours
70		seconds to minutes
80		fractions of seconds to seconds

4811 Each type of microorganism has a finite range of temperatures over which it can grow, some
4812 preferring lower temperatures, others higher temperatures. Note also that the effect of
4813 temperature depends on the temperature range considered. At low temperature, survival is
4814 enhanced, while at intermediate temperatures, growth rate increases with increased temperature.
4815 At temperatures above the limit for growth, however, death results at a rapidly increasing rate with
4816 increasing temperature.

4817 Each organism also has a finite range for growth as a function of pH, water activity, organic acid
4818 level, preservatives, etc., so that there are upper and lower limits for each factor, as well as an
4819 optimal level at which the growth rate is fastest. In general, the inhibitory effects of suboptimal
4820 factors interact both to reduce the range of each factor over which growth is possible when one or
4821 more factors are suboptimal, and to reduce the overall growth rate. At conditions beyond those that
4822 allow growth, stasis – or more probably death – will result at a rate dependent on the conditions but
4823 that is apparently most strongly influenced by the temperature (McQuestin, Shadbolt and Ross,
4824 2009; Zhang *et al.*, 2010).

4825 The growth of microorganisms in a unit of food follows the pattern of a 'batch' culture, often with a
4826 period of adjustment ('lag'), involving no growth, followed by exponential growth until some
4827 maximum population density (MPD) is reached and population growth ceases (see Figure 13). For
4828 many organisms and many foods, the MPD is in the range 10^9-10^{10} cells per gram, ml or cm^2 of food.

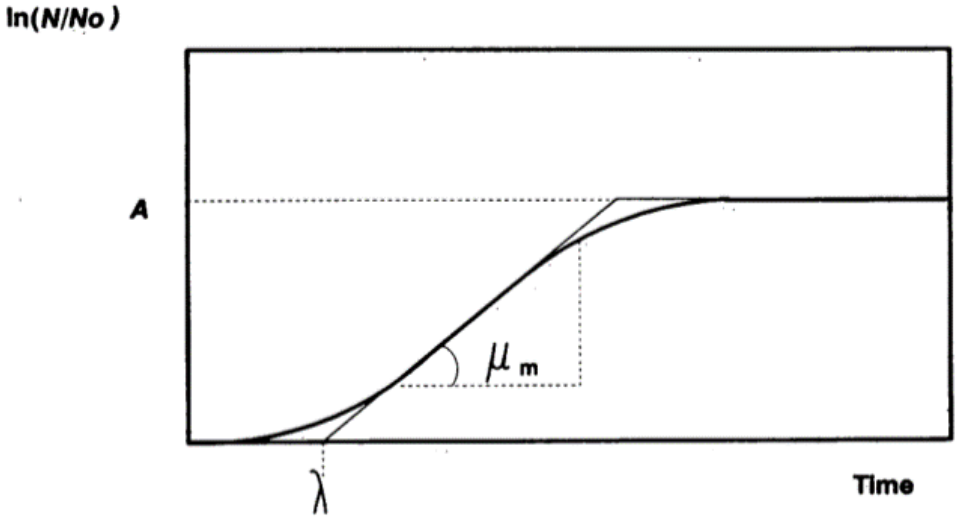

Figure 13: Example of a typical growth curve where A denotes the maximum population density, λ denotes the lag and μ_m denotes the maximum growth rate (Fig 1 of Zwietering et al., 1990)

Similarly, the death or inactivation of microorganisms in a unit of food is characterised by an initial period of no decrease in the microbial population ('shoulder'), followed by an exponential death phase until the 'tail' is reached and population decline ceases.

Although the ecology of microbiological hazards in food can be complex, predictive microbiology models can be used to estimate changes in microbial levels in foods as the product moves through the farm-to-fork chain. Ross (2008) provides a detailed discussion of the microbial ecology of foods in the context of the exposure assessment part of risk assessment.

12.1.2 *Predictive microbiology*

In recent years, significant advances have been made in the field of predictive microbiology. Some models are based on data obtained from liquid microbiological media and have been developed to predict the microbial behaviour when the physicochemical characteristics of the food (e.g. pH, water activity, organic acids concentrations) and the storage temperature are known. Some of these models can fail to accurately describe the microbial behaviour in foods, although the more robust models of this type have been validated in foods. Some other models have been developed to predict the behaviour of micro-organisms in particular foods whatever their storage conditions might be. The food-based models can effectively describe the impact of storage conditions on a specific food but their ability to describe the impact of the variability of physicochemical characteristics of the food or to make predictions in other foods is questionable. Some intermediate approaches have also been developed trying to overcome the limitations of these two major approaches. For certain products, it has been shown that proliferation (rate or extent, or both) of the spoilage microflora of a product influences the behaviour of the pathogen concerned, e.g. *L. monocytogenes* on cheese and cold-smoked salmon (Giménez and Dalgaard, 2004; Mellefont, McMeekin and Ross, 2008; Cadavez *et al.*, 2019).

For many bacterial pathogens, responses to environmental conditions have been described and summarized in mathematical models that can be used to predict their behaviour in foods, including growth rate, lag time, death rate, probability of growth occurring, and probability of toxin production within the storage life of the product. Models relating the number of a microbial organism and time, assuming that all other factors are constant, are known as primary models (Buchanan, 1993).

The physiological and physical state of the microorganism in the food remains a relatively unexplored area. Stress, injury and recovery also affect the initiation of growth. Spores will have a distribution of germination/outgrowth/lag times. Many studies use stationary phase cells grown in a nutrient-rich broth at favourable temperatures, and the predicted lag phase duration represents those conditions; cells that contaminate a food may be in a different physiological state. The extent to which the organism is clustered or aggregated may influence growth, survival and cross-contamination.

In predictive microbiology, foods are characterized in terms of their properties that most affect microbial growth and survival, such as temperature, pH, organic acid levels, salt levels and preservative levels. Microbial responses to analogous conditions are systematically studied and quantified, usually in a simplified laboratory broth model system under static and axenic conditions. The data are collated and summarized as predictive mathematical models. In particular, models that relate these properties to growth rate are known as secondary models (Buchanan, 1993).

Tertiary models are usually considered models that combine primary and secondary models (Buchanan, 1993). However, it has been argued that the term 'tertiary models' "should be used for patterns in the parameters of the secondary models as a function of the organism and the nutrient source" (Baranyi, Buss da Silva and Ellouze, 2017).

Conditions actually experienced by the foods and microbes are dynamic, and the effects of those conditions on rates of growth or inactivation have to be mathematically integrated over time for each of those distinct processes or stages. Thus, measurements of processing and handling parameters, and the duration for which these conditions are experienced, are integrated and used to predict changes in hazard levels (i.e. population size or concentration) in the food or water. Some predictive microbiology models, however, recreate the growth curve, i.e. the number of cells present, assuming a defined starting level, as a function of incubation time. Outputs from such models would normally have to be converted to rates of growth before their application in exposure assessment models.

A potential weakness of many predictive models is that they are developed in laboratory broth media, in which factors such as interactions with other microbes in the food or effects due to the physical structure of some foods are not observed. In general, these limitations relate to a few specific types of products, e.g. lactic acid bacteria may suppress pathogen growth in vacuum-packed or modified-atmosphere packed foods, matrix effects may be important in water in oil emulsions (e.g. butter). While most models have been developed in laboratory broth, models for some microorganisms have been developed in specific foods of concern or interest.

12.1.3 *Model types and modelling tools*

Models are available that describe:

- Rates of growth as a function of multiple environmental factors.
- Rates of inactivation, most as a function of a single lethal factor. One should be aware, however, that microbial inactivation is usually considered a stochastic process, i.e. the probability of survival of cells decreases (more or less) exponentially per unit of time. Thus, although the number of viable cells in an individual unit of food may be predicted to be less than one, one might still find survivors if a larger unit of the product (e.g. the total volume of a batch), or many units of the product, were examined or considered.

- Limits to growth as a function of multiple environmental factors, so-called 'growth/no growth' or 'interface models'. Absolute limits to growth of many pathogens due to individual environmental variables have been documented (ICMSF, 1996).
- Probability of growth or toxigenesis within a defined period as a function of multiple environmental factors.

In addition to numerous small-scale research projects to model microbial responses in foods, two large-scale predictive microbiology research programmes were undertaken in the early 1990s. They were funded by the governments of the United States of America and of the United Kingdom and resulted in the development of a suite of models for responses of populations of foodborne microbial pathogens and some spoilage organisms. The outcomes of those programmes, and subsequent developments, are now available without cost through the Predictive Microbiology Information Portal[17] which hosts the Pathogen Modelling Program and links to ComBase and ComBase Predictor[18]. These software packages include growth models for many pathogens and some spoilage organisms, and inactivation models for some pathogens. ComBase is a database of observations for many published and unpublished sources on microbial growth and inactivation rates, and at the time of writing contains approximately 60,000 records. The database is derived from the USA and UK government-funded research programmes referred to above, from data extracted from the published literature and from data (both published and unpublished) donated by researchers and research organizations around the world. Additional models for a range of pathogens and spoilage organisms are also available (Microsoft Windows only) from the Danish Technical University – Food Spoilage and Safety Predictor Web site[19]. Comprehensive lists of predictive microbiology modelling tools are available on the Combase[20] and the OpenML for Predictive Modelling in Food[21] websites, as well as Koutsamanis *et al.* (2016). The available tools offer a variety of utilities for the majority of foodborne pathogens including databases, fitting tools, predictions for growth, growth/no growth and inactivation, probabilistic models, and risk assessment modules. This allows for a wide range of applications including exposure assessment. The most important benefit for the users, however, is that software can assist decision-making in a short-time frame and allow practices to be actioned almost in real time.

Additionally, there are many modelling programmes and studies that have not resulted in the release of software but that are published (often including the data on which the model is based) in the scientific literature. These can be found readily by undertaking a literature search.

The integration of models for microbial growth, growth limits or inactivation into unified models that can predict both increases and decreases in microbial populations over time will also improve the utility of predictive models for exposure assessment. Several unified models have been proposed, but none have been widely used or endorsed.

Many reviews of predictive microbiology, including potential pitfalls, have been published. McMeekin *et al.* (1993) and Ross *et al.* (2014) provide a good introduction to the concept and its practical application, and the texts edited by McKellar and Lu (2003), Brul *et al.* (2007) and Pérez-Rodríguez and Valero (2013) provide more contemporary reviews of the state of the art. An

[17] https://portal.errc.ars.usda.gov/ accessed 20 June 2019
[18] http://www.combase.cc accessed 20 June 2019
[19] http://fssp.food.dtu.dk/ accessed 20 June 2019
[20] https://www.combase.cc/index.php/en/8-category-en-gb/21-tools accessed 29 November 2018
[21] https://sourceforge.net/p/microbialmodelingexchange/wiki/Tools/ accessed 29 November 2018

extensive listing of available predictive microbiology models was presented in Ross and Dalgaard (2004).

12.2 Application of predictive microbiology within exposure assessment

In practice, two features of a predictive microbiology model are critical to its utility. One is the ability to predict accurately microbial responses under all conditions to which the model applies. Evaluation of this ability is loosely termed 'model validation' (see Section 16.2.3). The second is the range of independent variables and variable combinations to which the model applies – if the model does not include terms for all factors of relevance to the microbial ecology of the hazard in the food, then that model is 'incomplete.' While predictive microbiology has matured considerably as a science discipline over the last two decades many currently available models are still incomplete or unvalidated, or both. Thus, exposure modelling should include consideration of the validity and reliability of predictive microbiology models, if used.

12.2.1 *Range of model applicability*

No predictive models currently in use are fully mechanistic (i.e. derived entirely from fundamental theoretical bases), therefore microbial growth or death cannot be reliably predicted in a food in which the conditions are beyond the range of any individual factor included in the data used to develop the model (i.e. predictions should be made by interpolation only).

Different models have different interpolation regions depending on the experimental design used to develop the model. The determination of the true interpolation region and the consequences of extrapolation were discussed by Baranyi *et al.* (1996). Those authors concluded that models that were over-fitted using a large number of parameters were more prone to unreliability resulting from inadvertent extrapolation, because the predictions of the model often changed dramatically near the limits of the interpolation region.

Inadvertent extrapolation can also occur when using stochastic modelling techniques to describe effects of fluctuating variables. This problem may occur for any factor, but temperature is the factor most likely to fluctuate in most real-world examples. Consideration should be given to truncating the tails of the temperature (and other) distributions used to predict microbial growth or death, if necessary, to match the interpolation range of the predictive microbiology model used. This should be done by utilizing a suitable 'truncated' distribution so that the mean, variance and other properties of the chosen distribution are not changed in unintended ways (Johnson, A. W. Kemp and Kotz, 2008). The growth limits for the pathogen of concern, and potential for inactivation (if conditions are beyond those limits) should be considered and included in exposure modelling. Growth/no growth models may assist in this regard and have been included in some exposure assessment models.

12.2.2 *Spoilage microbiota*

The effect of spoilage bacteria on the shelf life of the product should also be considered. Conditions that lead to rapid growth of pathogens may also lead to rapid microbial spoilage. Contaminated products that are obviously spoiled are less likely to be consumed, and thus do not lead to foodborne disease, despite that fact that they contain a microbiological hazard. Thus, it may be necessary to consider the effect of storage conditions on the shelf life of the product in case unrealistically long times at high temperatures are simulated. This can be implemented by correlating model variables that affect growth (e.g. storage time and temperature). Stochastic modelling texts offer advice on how such correlations can be included in models and examples include Ross *et al.* (2009), Smith *et al.* (2013) and Kiermeier *et al.* (2015).

On a related topic, other microorganisms growing in the food can influence the potential growth of pathogens. Exposure assessments that rely on empirical data derived from pure culture broth systems are likely to overestimate potential growth of pathogens in food matrices due to the co-existence of numerous competing bacterial population (Coleman, Sandberg and Anderson, 2003). Pathogen growth rates and maximum densities are thought to be a function of the total microbial community composition and density in the food due to competition for nutrients, the production of inhibitory substances, and overall density (Powell, Schlosser and Ebel, 2004). The final cell density of a pathogenic bacterium can be suppressed when the total concentration of all bacteria in the food reaches stationary phase, a phenomenon that has been termed the 'Jameson Effect' (Jameson, 1962; Stephens *et al.*, 1997) and reported by many authors (e.g. Ross, Dalgaard and Tienungoon, 2000; Le Marc, Valík and Medveďová, 2009; Al-Zeyara, Jarvis and Mackey, 2011). In many foods, this effect will not happen before spoilage occurs, but in vacuum-packed or modified-atmosphere packed foods such as processed meats and lightly preserved fish, lactic acid bacteria can reach stationary phase without causing overt spoilage and limit the growth of pathogens to safe levels within the acceptable shelf life of the product.

12.2.3 *Sources of variability and uncertainty*

In stochastic modelling, it is important to characterize the magnitude of the variability and its distribution about the mean. Traditionally, the approach to fit predictive microbiological models was through a 2-step fitting approach. For example, first primary models were fitted separately for each temperature and the model parameter estimates were extracted (esp. max growth rate). Subsequently these estimates were used as the response for the secondary model, i.e. the model was fitted to relate them to temperature; the implicit assumption is that the parameter estimates are known values rather than estimates. This approach was likely due to the nonlinear nature of problem (i.e. when primary and secondary models are combined) and the result of limited computing power in the early days of the discipline. However, fitting nonlinear models is no longer a major problem, though the actual fitting process can still be problematic; good starting estimates and suitable parameter transformations can help in this regard. In addition, it has been shown that the 1-step model fitting process, i.e. where the primary and secondary models are combined and estimated in a single model, is more efficient, and hence to be preferred, than the 2-step process (Jewell, 2012; Dolan and Mishra, 2013; Huang, 2017).

Distribution of response times

Using the limited amount of replicated published data concerning growth rate estimates under varying environmental conditions, Ratkowsky *et al.* (1991) concluded that growth rates became increasingly variable at slower growth rates. Microbial response times or rates as a function of environmental conditions are often not normally distributed. Distributions describing growth rate and/or response time variability in response to temperature have been described by various researchers (Ratkowsky *et al.*, 1991, 1996; Alber and Schaffner, 1992; Dalgaard *et al.*, 1994; Zwietering *et al.*, 1994). Ratkowsky (1992) presented a general relationship between the variance in growth response times and the mean of those responses for a range of possible distribution types.

Sources and magnitude of errors

Model predictions can never perfectly match observations or represent reality. Each step in the model construction process introduces some error as described below (Cullen and Frey, 1999; Ross, McMeekin and Baranyi, 2014).

- *Homogeneity* error arises because some foods are clearly not homogeneous. Current predictive models do not account for this non-homogeneity of foods.

- *Completeness* error in predictive models arises because the model is a simplification, and other food effects and microbial ecology effects (structure, competition, etc.) that are difficult to quantify are not included in currently available models.
- *Model function* error is similar to completeness error and arises mainly from the compromise made when using empirical models, namely that the model is only an approximation to reality.
- *Measurement* error originates from inaccuracy in the limitations in the measurement methods used to collect raw data that are used to estimate the parameters of a model.
- *Numerical procedure* error includes all errors arising from procedures used for model fitting and evaluation, some of which are only methods of approximation.

Like all statistical models, the fit of the model should be checked graphically against the actual observations. Sometimes the fitted model clearly doesn't match the data very well, in which case a different model formulation may need to be considered.

As rule of thumb, when constructing a predictive microbiology model from data, each additional variable increases the error in the estimate of the specific growth rate by approximately 10% (Ross, McMeekin and Baranyi, 2014). In other words, confidence in the predicted growth rate, and total predicted growth declines when more variables that affect the growth rate are considered. The significance of this for predicted exposure depends on the amount of growth predicted to occur. For a three-factor/variable model the magnitude of the 'error' in terms of growth rate and log number of cells would be around ±30%, irrespective of the amount of growth predicted. However, in many situations, probability of infection (and thus risk) is related to the absolute number of cells ingested, not the logarithm of dose. Thus, if one generation of growth (0.30 $\log_{10}$) were predicted (assuming the lag time and maximum population density are known exactly and not estimated), the error in the predicted number of cells would be ±(0.30 $\log_{10}$ × 0.3), i.e. ±23% of the estimate. If 10 generations of growth were predicted, the 'error' would be ±(3.00 $\log_{10}$ × 0.3) which, in terms of numbers of cells would be ±800%. If lag time and MPD are also estimated, then these errors will be larger.

13. Dose-Response

The assumptions on which current models are based, their use and possible limitations are carefully considered in the following sections.

Dose-response modelling requires a combination of mathematics, statistics, human biology (infection process, immune system), microbiology and epidemiology. Different approaches are available for model fitting and the assumptions underlying the dose-response model need to be understood, assessed and reported/communicated.

The focus of these sections is on infectious and toxico-infectious hazards, as this has been the area of most development. However, it should be noted that this chapter provides an overview of dose-response models and the interested reader is directed to the review by Haas (2015), which provides information on D-R models not only for foodborne hazards.

13.1 The infectious disease processes

The biological basis for dose-response models derives from major steps in the disease process as they result from the interactions between the hazard, the host and the matrix. Figure 14 illustrates the major steps in the overall process, with each step being composed of many biological events. Colonization, toxin production, infection and illness can be seen as resulting from the hazard successfully passing multiple barriers in the host. These barriers are not all equally effective in eliminating or inactivating hazards and may have a range of effects, depending on the hazard and the individual. Each individual hazard has some particular probability to overcome a barrier, which is conditional on the previous step(s) being completed successfully, similar to the *hurdle concept* in food processing. The disease process as a whole, and each of the component steps, may vary by hazard and by host. Hazards and hosts can be grouped with regard to one or more components, but this should be done cautiously and transparently.

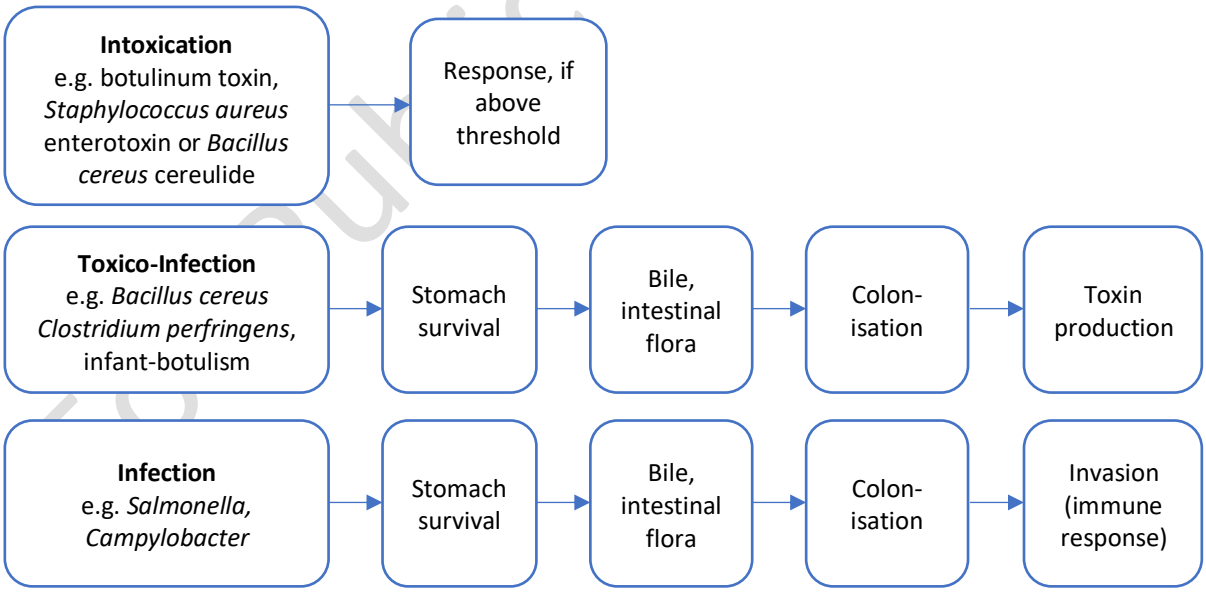

Figure 14: The major steps in the foodborne infectious disease process.

13.1.1 *Infection*

Infection is usually measured as a quantal response, i.e. the presence or absence of infection by some criterion. The use of continuous-response variables, e.g. an antibody titre, may be useful for further development of dose-response models.

There are usually many different and simultaneous signs and symptoms of illness in any individual, and the severity of symptoms varies among hazards, strains and among hosts infected with the same hazard. The extent of illness is therefore a process that can also be measured on a multidimensional, quantitative, continuous scale (number of stools passed per day, body temperature, laboratory measurements, etc.).

A wide variety of case definitions for gastrointestinal illness are used in the literature, based on a variable list of symptoms, with or without a specified time window, and sometimes including laboratory confirmation of etiological agents. This lack of standardization severely hampers integration of data from different sources.

13.1.2 *Sequelae and mortality*

In a small fraction of ill persons, chronic infection or sequelae may occur. Some pathogens, such as *Salmonella enterica* serotype Typhi, are invasive and may cause bacteraemia and systemic infections. Other pathogens produce toxins that may result not only in enteric disease but also in severe damage in susceptible organs. An example is haemolytic uraemic syndrome, caused by damage to the kidneys from Shiga-like toxins of some *Escherichia coli* strains. Complications may also arise by immune-mediated reactions: the immune response to the pathogen is then also directed against the host tissues. Reactive arthritis, including Reiter's syndrome, and Guillain-Barré syndrome are well known examples of such sequelae. The complications from gastroenteritis normally require medical care, and frequently result in hospitalization. There may also be a risk of mortality in relation to sequelae, and not all patients may recover fully, but may suffer from residual symptoms, which may last a lifetime. Therefore, despite the low probability of complications, the public health burden may be significant. Also, there is a direct risk of mortality related to acute disease, in particular in the elderly, neonates and severely immunocompromised.

In the context of a risk assessment, the number of cases with sequelae and complications are usually ascertained on a proportional basis, similar to the approach used by the WHO Foodborne Disease Burden Epidemiology Reference Group (Section 4.2, WHO, 2015).

13.2 Modelling concepts

13.2.1 *The particulate nature of the inoculum.*

It is commonly assumed that the organisms are randomly distributed in the inoculum, but this is rarely the case. The Poisson distribution is generally used to characterize the variability of the individual doses when pathogens are randomly distributed.

Compound distribution or over-dispersion may result from two different mechanisms:

a) A "unit" as detected by the measurement process (e.g. a colony-forming unit (CFU), a tissue culture infectious dose, or a Polymerase Chain Reaction (PCR) detectable unit) may, due to aggregation, consist of more than one particle. This is commonly observed for viruses (e.g. Teunis *et al.*, 2008), but may also be the case for other pathogens (e.g. Jongenburger *et al.*, 2011). The degree of aggregation strongly depends on the methods used for preparing the inoculum. It is important to know whether the aggregates remain intact during inoculum preparation or in the gastrointestinal tract.
b) In a well-homogenized liquid suspension, single disaggregated organisms will be more or less randomly distributed. If the inoculum consists of a solid or semisolid food matrix, however, spatial clustering may occur and result in over-dispersion of the inoculum (e.g. Jongenburger *et al.*, 2012). This aspect of spatial clustering may differ between the data underlying the dose-response model and the actual exposure scenario.

The reason why knowing about aggregation is important is that it can have an effect on the dose-response model and thus the estimated 50% infectious dose ID_{50}. For example, for norovirus it was found that the ID_{50} was 1,015 genome copies for the aggregated inoculum, while for the disaggregated virus the ID_{50} was only 18 viruses (Teunis et al., 2008) – approximately two orders of magnitude lower!

13.3 Selection of models

Specific properties in the data become meaningful only within the context of a model. Different models may, however, lead to different interpretations of the same data, and so a rational basis for model selection is needed. Different criteria may be applied when selecting mathematical models. For any model to be acceptable, it should satisfy the statistical criteria for goodness of fit, in particular, residual plots are essential tools for assessing goodness of fit. In the case of more than one model fitting equally well, goodness of fit statistics, such as the various likelihood-based Information Criteria, can be used to select "the best" (Dziak *et al.*, 2018). However, many different models will usually fit a given data set (e.g. Holcomb *et al.*, 1999) especially due to the large variability and uncertainty in the data and therefore goodness of fit is not a sufficient criterion for model selection. Additional criteria that might be used are conservativeness, flexibility, parsimony and biological plausibility.

A conservative model is one that tends to over-predict the response of interest (i.e. in the context of a D-R model this is the probability of infection or illness). However, conservativeness can be approached in different ways: "Is the model structure conservative?", "Are parameter estimates conservative?", "Are specific properties, e.g. prediction at low doses, of the model conservative?" and so forth. It is not recommended to build conservativeness into the model structure itself.

From a risk assessment perspective, a model should be restricted to describing the data and trying to discriminate the biological signal from the noise. Adding parameters usually improves the goodness of fit of a model but using a flexible model with many parameters may result in overfitting (Lever, Krzywinski and Altman, 2016; Steyerberg *et al.*, 2010) – a lack of parsimony – and greater uncertainty of estimates, especially for extrapolated doses.

It is recommended that dose-response models be biologically plausible. For example, a quadratic model may fit a given data set well, or even better than an alternative model, yet the quadratic model is not biologically plausible and will result in inappropriate predictions when extrapolated to very small or large doses. Note that it is generally not possible to "work back", i.e. to deduce the assumptions underlying a given model formula. There is a problem of identifiability: the same functional form may result from different assumptions, while two (or more) different functional forms (based on different assumptions) may describe the same dose-response data equally well. This can result either in very different fitted curves if the data contain little information, or virtually the same curves if the data contain strong information. However, even in the latter case, the model extrapolation may be very different. This means that a choice between different models or assumptions cannot be made on the basis of data alone (e.g. FAO/WHO, 2011b, Annex A1.1.1).

13.3.1 *Dose-infection models*

Consider a host who ingests exactly one cell of a pathogenic microorganism. According to the single-hit hypothesis, the probability that this pathogen will survive all barriers and colonize the host has a non-zero value of p. Thus, the probability of the host not being infected is $1-p$. If a second cell of the pathogen is ingested, and the hypothesis of independent action is valid, then the probability of the host not being infected is $(1-p)^2$. For n pathogens, the probability of not being infected is $(1-p)^n$. Hence, the probability of infection of a host that ingests exactly n pathogens can be expressed as:

$$P_{inf}(n \mid p) = 1 - (1 - p)^n$$

When the discrete nature of pathogens is also taken into account, these concepts lead to the single-hit family of models.

13.3.2 *Dose-illness models*

The default assumption of constant probability models for illness given infection leads to the conclusion that the only difference between dose-infection and dose-illness models is that the dose-illness models do not need to reach an asymptote of 1 because the probability of illness given infection can be something less than one when the probability of illness given dose approaches 1 (Teunis and Havelaar, 2000; Teunis, Nagelkerke and Haas, 1999). As such they essentially still belong to the family of hit-theory models.

13.3.3 *Sequelae and mortality*

Given illness, the probability of sequelae or mortality, or both, depends on the characteristics of the pathogen, but more importantly on the characteristics of the host. Sequelae or mortality are usually rare events that affect specific subpopulations. These may be identified by factors such as age or immune status, but increasingly genetic factors are being recognized as important determinants. As for dose-illness models , the current possibilities are mainly restricted to constant probability models (e.g. FSIS, 2001). In the case of mortality, the proportion of infected patients who died is known as the *mortality ratio*. Stratification appears to be necessary in almost all cases where an acceptable description of risk grouping is available.

13.4 Extrapolation

13.4.1 *Low dose extrapolation*

Dose-response information is usually obtained in the range where the probability of observable effects is relatively high. In experimental studies using human or animal subjects, this is related to financial, ethical and logistical restrictions on group size. In observational studies, such as outbreak studies, low dose effects can potentially be observed directly, but in these studies only major effects can be distinguished from background variation. The single-hit family of models (further exemplified in Section 13.5) is characterized by linear low dose extrapolations on the log/log scale, or even on the arithmetic scale. That is, in the low dose range, the probability of infection or illness increases linearly with the dose and hence on the log-scale, these models have a slope of 1 at low doses (see for example Figure 8). Some model examples include:

• The Binomial model	$P_{inf}(n \mid p_1) = 1 - (1 - p_1)^n$	$P_1 = p_1$
• The linear model	$P = r \times D$	$P_1 = r$
• The exponential model	$P = 1\text{-}\exp(-r \times D)$	$P_1 = 1\text{-}\exp(-r) \approx r$
• Beta-Poisson model	$P = 1\text{-}[1+D/\beta]^{-\alpha}$	$P_1 \approx (\alpha/\beta)$
• The hypergeometric model	$P = 1\text{-}{}_1F_1(\alpha, \alpha+\beta, -D)$	$P_1 \approx \{\alpha/(\alpha+\beta)\}$

where D = mean ingested dose and r, α and β are model parameters. Note that if α > β, the probability of infection predicted by the Beta-Poisson model is larger than one, which is not biologically plausible.

13.4.2 *Extrapolation in the pathogen-host-matrix triangle*

Experimental datasets are usually obtained under carefully controlled conditions (e.g. using specific strains), and the data apply to a specific combination of pathogen, host and matrix. In actual exposure situations, there is more variability in each of these factors, and dose-response models need to be generalized. Assessing such variability requires the use of multiple datasets that capture

5219 the diversity of human populations, pathogen strains and matrices. Failure to take such variation
5220 into account may lead to underestimation or overestimation of the actual risk of the outcome of
5221 interest.

5222 When developing dose-response models from multiple datasets, one should use all the pertinent
5223 data. This requires that the risk assessors make choices about how to use different datasets. Such
5224 choices should be based on objective scientific arguments but will inevitably include subjective
5225 arguments. Such arguments should be fully and transparently documented and ideally be discussed
5226 with the risk manager and their significance and impact for risk management considered. The
5227 credibility of dose-response models increases significantly if dose-response relations derived from
5228 different data sources are consistent.

5229 When combining data from different sources, a common scale on both axes is needed. This often
5230 requires adjusting the reported data to make them comparable. For the dose, test accuracy, sample
5231 size, etc., need to be taken into account. For the response, a consistent case definition is needed, or
5232 the reported response needs to be adjusted to a common denominator (e.g. infection × conditional
5233 probability of illness given infection). Combining data from different sources within a single
5234 (multilevel) dose-response model requires thorough statistical skills and detailed insight into the
5235 biological processes that generated the data. An example is the multilevel dose-response model that
5236 has been developed for different isolates of *Cryptosporidium parvum* (Teunis, Chappell and
5237 Okhuysen, 2002a). The issue of combining data from different outbreak studies is discussed in the
5238 FAO/WHO risk assessments of *Salmonella* in eggs and broiler chickens (FAO/WHO, 2002a).

5239 Dose-response relations where the hazard only affects a portion of the population do require that
5240 subpopulations be separated from the general population to generate meaningful results. Using
5241 such stratified dose-response models in actual risk assessment studies requires that the percentage
5242 of the population that is actually susceptible can be estimated. Consideration of such subpopulations
5243 appears to be particularly important when attempting to develop dose-response relations for
5244 serious infections or mortality. However, it would also be pertinent when considering a hazard for
5245 which only a portion of the population can become infected, e.g. not all people are susceptible to
5246 norovirus infection (Teunis *et al.*, 2008).

5247 A particular and highly relevant aspect of microbial dose-response models is the development of
5248 specific immunity in the host. Most volunteer experiments have been conducted with test subjects
5249 selected for absence of any previous contact with the pathogen, usually demonstrated by absence of
5250 specific antibodies. The actual population exposed to foodborne and waterborne pathogens will
5251 usually be a mixture of totally naive persons and persons with varying degrees of protective
5252 immunity. No general statements can be made on the impact of these factors. This is strongly
5253 dependent on the pathogen and the host population. Some pathogens, such as many childhood
5254 diseases and the hepatitis A virus, will confer lifelong immunity upon first infection whether clinical
5255 or subclinical, whereas immunity to other pathogens may wane within a few months to a few years,
5256 or may be evaded by antigenic drift. At the same time, exposure to non-pathogenic strains may also
5257 protect against virulent variants. This principle is the basis for vaccination, but has also been
5258 demonstrated for natural exposure, e.g. to non-pathogenic strains of *Listeria monocytogenes*
5259 (Notermans *et al.*, 1998). The degree to which the population is protected by immunity depends to a
5260 large extent on the general hygienic situation. In many developing countries, large parts of the
5261 population have built up high levels of immunity, and this is thought to be responsible for lower
5262 incidence or less serious forms of illness. Some examples are the predominantly watery form of
5263 diarrhoea by *Campylobacter* spp. infections in children and the lack of illness from this organism in
5264 young adults in developing countries. The apparent lack of *E. coli* O157:H7-related illness in Mexico

has been explained as the result of cross-immunity following infections with other *E. coli*, such as enteropathogenic *E. coli* strains that are common there. Obviously, age is an important factor in this respect, as older people will have greater likelihood of prior exposure than children. In contrast, in the industrialized world, contact with enteropathogens is less frequent and a larger part of the population is susceptible. This also highlights that dose-response models may not be globally applicable.

Incorporating the effect of immunity in dose-response models has received little attention. The absence of accounting for immunity in dose-response models may complicate interpretations, and comparisons among geographic regions. This is particularly likely to be a problem with common infections such as *Campylobacter* spp., *Salmonella* spp. and pathogenic *E. coli*. Immunity may affect the probability of infection, the probability of illness given infection, or the severity of illness. There are currently only few data sets available on which to base model development. Where such data are available, a simple and possibly effective option would be to resort to stratified analysis and divide the population into groups with different susceptibility (e.g. FDA/FSIS, 2003; Pouillot *et al.*, 2015b; Teunis *et al.*, 2008). Experimental work on infection of volunteers having different levels of acquired immunity to *Cryptosporidium parvum* was analysed with a dose-response model that includes the effects of immunity (Messner and Berger, 2016; Teunis, Chappell and Okhuysen, 2002b).

Stratified analysis can also be useful when dealing with seemingly outlying results, which may actually indicate a subpopulation with a different response. Removal of one or more outliers corresponds to removing (or separately analysing) the complete group from which the outlying results originated. Where a specific reason for the separation cannot be identified, there should be a bias toward being inclusive in relation to the data considered. As for all data analysis, any exclusion of the outlying data should be scientifically justified and clearly communicated to ensure the transparency of the assessment.

13.5 Dose-response model fitting approaches

According to the single-hit hypothesis (see Section 6.3 and 13.3.1), the probability of infection of a host that ingests exactly *n* pathogens can be expressed as:

$$P_{inf}(n \mid p) = 1 - (1 - p)^n$$

This model is also called the *binomial dose-response model*. Starting from this basic function and taking the discrete nature of pathogens into account, a broad family of dose-response models (hit-theory models) can be derived. The most frequently used models are the exponential and the Beta-Poisson models, which are based on further assumptions on the distribution of pathogens in the inoculum, and on the value of p. When the distribution of the organisms in the inoculum is assumed to be random, and characterized by a Poisson distribution, it can be shown (Teunis and Havelaar, 2000) that the probability of infection as a function of the dose is given by:

$$P_{inf}(D \mid p) = 1 - \exp\{-Dp\}$$

where D is the *mean* ingested dose (while the *n* above is the *exact* number of organisms ingested). This model gives virtually the same outcome as the above binomial model. If p is assumed to have a constant value *r* for any given host and any given pathogen, the simple exponential model results:

$$P_{inf}(D \mid r) = 1 - \exp\{-rD\}$$

When the dose is low and $rD << 1$, then this formula is approximated by a straight line, i.e.

5307 $P_{inf}(D \mid r) \approx rD$

5308 If the probability of starting an infection differs for any organism in any host, and is assumed to
5309 follow a beta-distribution, then:

5310 $P_{inf}(D \mid \alpha, \beta) = 1 - {}_1F_1(\alpha, \alpha+\beta, -D)$

5311 Where ${}_1F_1()$ is the Kummer confluent hypergeometric function (Abramowitz and Stegun, 1972),
5312 which can also be found in the Digital Library of Mathematical Functions (https://dlmf.nist.gov/). For
5313 $\alpha << \beta$ and $\beta >> 1$, P_{inf} is approximately equal to the Beta-Poisson formula:

5314 $P_{inf}(D \mid \alpha, \beta) \approx 1 - (1 + D/\beta)^{-\alpha}$

5315 As for the exponential model, when the dose is low and $D\alpha << \beta$, this formula is approximated by a
5316 straight line (which also holds for the exact form involving ${}_1F_1()$), i.e.

5317 $P_{inf}(D \mid \alpha, \beta) \approx D\,\alpha / \beta.$

5318 For both $\alpha \to \infty$ and $\beta \to \infty$, while $\alpha/\beta \to r$, the Beta-Poisson formula converts into the exponential
5319 model.

5320 Other assumptions for *n* or p_m lead to other models. For example, spatial clustering of cells in the
5321 inoculum can be represented by a negative binomial distribution or any other contagious
5322 distribution. However, this has little effect on the shape of the dose-response relationship (Haas,
5323 Rose and Gerba, 2014) although the limiting curve for the confidence interval is affected (Teunis and
5324 Havelaar, 2000). It is also possible to model p as a function of covariables, such as immune status or
5325 age.

5326 Using these models, it is possible to determine the dose below which the dose-response relationship
5327 is linear (Williams, Ebel and Vose, 2011a). If the exposure distribution is such that doses will be
5328 below this value, then the risk characterization is greatly simplified.

14. Uncertainty / Variability

Variability and uncertainty are frequently confused because both can be described by distributions. However, they have distinct meanings (Haas, Rose and Gerba, 2014; Nauta, 2000; Vose, 2008), and a common understanding between the risk manager and risk assessor of these concepts can greatly help in the risk assessment process. These topics are considered below.

14.1 Variability

Variability, also sometimes referred to as inter-individual variability, refers to real differences in values of some property of the individuals from a 'population' over time or space. The population could refer to people, units of food, a species of foodborne pathogen, etc. Examples of variable factors relevant to microbiological risk assessment include (but are not limited to) the storage temperatures of food products, seasonality of different food preparation methods (e.g. barbecuing), culinary practice, susceptibility to infection across subpopulations, consumption patterns across a region, differences in growth and inactivation characteristics and in virulence between strains, and product handling processes across different producers.

In some cases, some of the variability in the population can be explained by observable individual attributes or explanatory factors. For example, while the human population is heterogeneous, there may be discernible differences between identifiable subpopulations because they are for some reason less frequently exposed, or less susceptible, to the hazard of interest. Or there could be different methods of storing a food product, e.g. frozen, chilled and not chilled, leading to different potential for microbiological growth; the fractions of the food item that are stored in each manner need to be known or estimated, and they may vary over time.

Hence, variability is inherent in the population being studied and describes by how much a specific attribute differs between the units in that population. As a result, variability cannot generally be reduced by more accurate measurement or more data points, it can only be estimated more precisely. However, some sources of variability may be explained by having more information, such as, knowing whether a food product was stored frozen, chilled or not chilled.

In principle, variability can be described by listing the different values that the attribute can take. Often however, there are such a large number of values that it is more convenient to describe the variation using a probability distribution. For example, if it is considered an animal shedding an enteric bacterial pathogen, then there are only two possible values, that is, the animal is shedding, or it is not. In contrast, if it is considered the number of bacterial cells in a 10 g faecal sample, then the possible values are 0, 1, 2, 3, etc. Instead of enumerating all possible values, and the probability with which these outcomes can occur, it is usually preferred (if possible) to describe the possible outcomes by a mathematical distribution, such as the Poisson or the negative-binomial distributions. The use of some mathematical distributions is quite well established for some circumstances, for example, the binomial distribution is usually used to describe the number of infected animals sampled from a large herd (or alternatively the hypergeometric distribution can be used for small herds). Similarly, the concentration of microbial cells in a sample is often assumed to follow a log-normal distribution, although others may be more appropriate (e.g. Bassett *et al.*, 2010; Haas, Rose and Gerba, 2014; Vose, 2008). Where possible, the fit of the mathematical distribution used to model a particular situation should be checked against empirical data. Tools for this include:

- Density histograms with fitted distributions overlaid;
- Cumulative distribution plots with fitted distributions overlaid;
- Quantile-quantile plots; and
- Skewness-kurtosis plot (Cullen and Frey, 1999).

When there are discernible differences due to known factors, 'stratification' can be a practical method of addressing the population variability by recognizing those populations as discrete within the risk assessment. The properties of each subpopulation, or stratum, may still be described as a variable quantity, but with a different mean and spread of values. There are many ways of stratifying a human population such as demographic, cultural and other variables, but in microbial risk assessment stratifications are usually done in one of two ways. One is based on differences in exposure and the other is due to differences in susceptibility, usually related to well recognised subpopulations such as the very young, old, pregnant and immune-compromised (YOPI). Exposure and sub-population strata may be combined, that is, within the population of interest, if there is evidence of differences in susceptibility or differential exposure patterns, then consideration should be given to stratifying the risk accordingly.

These ideas are illustrated in Figure 15. Here, it is assumed that exposure depends on season (A and B) and producer (1 and 2), leading to 4 different distributions of exposure (A1, A2, B1, B2). In addition, it is assumed that there are two subpopulations, each of which has its own dose-response curve. The figure shows how each exposure model is combined with the appropriate dose-response model if exposure and dose-response are stratified in this way.

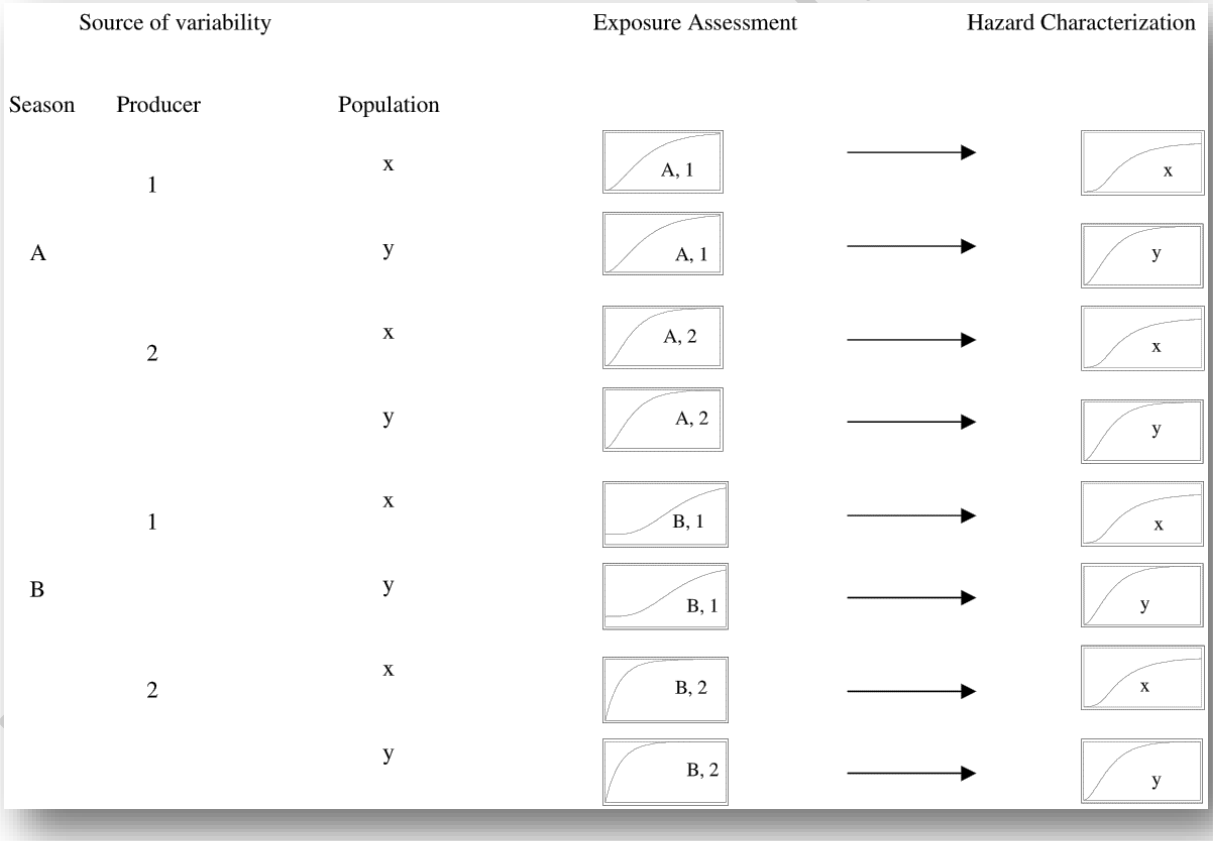

Figure 15: Linkage between exposure assessment and hazard characterization

With respect to qualitative and semi-quantitative risk assessments, one option for the inclusion of variability is to consider a number of scenarios that reflect the variability, e.g. near-optimal condition, normal situation and one or more adverse conditions. The risk assessment then evaluates each as a separately measured risk scenario, and the results are compared. The overall assessment of variability (and also uncertainty) will be evaluated in narrative terms such as 'small', 'small', etc.

This approach will make the effects of variability on the risk estimate more transparent. However, if the scenarios vary greatly in risk outcome, such an analysis may provide insufficient support for decision-making in the absence of any description of the relative likelihood of each scenario. It should be noted that risk can be dominated by, or at least strongly influenced by, the more extreme scenarios, e.g. conditions leading to relatively high risk, despite their lower probability. It is important that the risk assessor identifies the likelihood with which such scenarios could occur.

14.2 Uncertainty

Uncertainty arises due to a lack of knowledge and is sometimes termed epistemic uncertainty, lack-of-knowledge uncertainty, or subjective uncertainty. It is often stated that variability is a property of the system being studied, whereas uncertainty is a property of the methodology and data used. Assessments with different methodologies and data, etc. will have different levels of uncertainty regarding their outputs. An understanding of uncertainty is important because it provides insight into how the lack of knowledge can influence decisions. In the EFSA opinion on the principles and methods behind EFSA's Guidance on Uncertainty Analysis in Scientific Assessment (EFSA, 2018a) uncertainty is used as a general term referring to all types of limitations in available knowledge that affect the range and probability of possible answers to an assessment question. Available knowledge refers to the knowledge (evidence, data, etc.) available to assessors at the time the assessment is conducted and within the time and resources agreed for the assessment. When the uncertainty is large enough that there is ambiguity as to which risk management decision is preferred, then there may be value in collecting additional data or conducting additional research to reduce the uncertainty; it is the risk managers' role to decide if the uncertainty of a risk assessment output allows for a decision to be made or not. These aspects apply equally to all parts of the qualitative-to-quantitative continuum of risk assessment.

In contrast to variability, uncertainty is not inherent in the population, but a result of limited information and our lack of knowledge. Consequently, well targeted collection of data or information can usually help reduce uncertainty. For example, the uncertainty in the parameter estimates from a linear regression model can be reduced when more data from the same population can be incorporated into the model fit. Similarly, uncertainty in the processing practices used to manufacture a food product can be reduced by visiting different manufacturing facilities (of different sizes) to gain a better of what actually happens in practice.

Uncertainty is associated not only with the inputs to an assessment model, but also regarding the scenarios assumed for the assessment and the model itself. Sources of scenario uncertainty include potential misspecification of the harmful agents of concern, exposure pathways and vectors, exposed populations, and the spatial and temporal dimensions of the problem.

Sources of model uncertainty include model structure, detail, resolution, validation or lack thereof, extrapolation, and boundaries of what is included and what is excluded from the model. A list of most common types of uncertainty affecting scientific assessments associated with the inputs and the methodology was identified by EFSA (2018a) and these are presented in Table 40. In addition, Morgan and Henrion (1992) and Cullen and Frey (1999) provide examples of sources of uncertainty in risk assessment.

Table 40: List of most common types of uncertainty affecting risk assessments associated with the inputs and the methodology (EFSA, 2018a).

Uncertainties associated with assessment inputs	Uncertainties associated with assessment methodology
Ambiguity	Ambiguity

Accuracy and precision of the measures	Excluded factors
Sampling uncertainty	Distributional assumptions
Missing data within studies	Use of fixed values
Missing studies	Relationship between parts of the assessment
Assumptions about inputs	Evidence for the structure of the assessment
Statistical estimates	Uncertainties relation to the process for dealing with evidence from the literature
Extrapolation uncertainty (i.e. limitations in external validity)	Expert judgement
Other uncertainties	Calibration or validation with independent data
	Dependency between sources of uncertainty
	Other uncertainties

14.3 Uncertainty Analysis

Uncertainty analysis is the process of identifying limitations in scientific knowledge and evaluating their implications for scientific conclusions (EFSA, 2018b). It is therefore relevant in all risk assessments to ensure that the conclusions provide the risk managers reliable information for decision making. The form and extent of uncertainty analysis, and how the conclusions should be reported, vary widely depending on the nature and context of the assessment and the degree of uncertainty that is present.

In a Guidance on Uncertainty Analysis (EFSA, 2018a) EFSA presented the main elements of an uncertainty analysis as the following:

1. Identifying uncertainties affecting the assessment.
2. Prioritising uncertainties within the assessment
3. Dividing the uncertainty analysis into parts.
4. Ensuring the questions or quantities of interest are well-defined.
5. Characterising uncertainty for parts of the uncertainty analysis.
6. Combining uncertainty from different parts of the uncertainty analysis.
7. Characterising overall uncertainty.
8. Prioritising uncertainties for future investigation.
9. Reporting uncertainty analysis.

Identifying the various uncertainties affecting the risk assessment outputs is necessary in every assessment and should be done in a structured way to minimise the chance of overlooking relevant uncertainties. Although it is often efficient to concentrate detailed analysis on the most important sources of uncertainty, the identification of uncertainties needs to be as comprehensive as possible. Risk assessors should examine in a systematic way every part of their assessment in order to identify all uncertainties, including those related to the inputs of the assessment as well as the methods used in the assessment (see Table 40 above).

Prioritising uncertainties within the risk assessment plays an important role in planning the uncertainty analysis, enabling the assessor to focus detailed analysis on the most important uncertainties and address others collectively when evaluating overall uncertainty. Prioritisation can be done by expert judgement during the planning process. In more complex risk assessments

uncertainties can be prioritized explicitly using sensitivity analysis (see Chapter 15). Depending on the methods and data used, it may be sufficient to characterise overall uncertainty for the whole assessment directly, by expert judgement. In other cases, it may be preferable to evaluate uncertainty for some or all parts of the assessment separately and then combine them to evaluate the overall uncertainty, either by calculation or expert judgement.

Each parameter of interest must be well-defined. This is necessary to ensure the parameter can be estimated appropriately and to make it possible to express uncertainty clearly and unambiguously.

Sometimes risk assessors choose or need to divide the uncertainty analysis into parts. In these cases, there may be a need to combine the different parts of the uncertainty analysis if an overall estimate of uncertainty is needed.

The element of overall uncertainty characterization includes the quantitative expression of the overall effect of as many as possible of the identified uncertainties on the conclusions and the qualitative description of any uncertainties that remain unquantified. In assessments where the impact of one or more uncertainties cannot be characterised, it must be reported that this is the case and that conclusions are conditional on assumptions about those uncertainties; these assumptions also need to be specified.

Prioritising uncertainties for future investigation is implicit or explicit in any assessment where recommendations are made for future data collection or research; these priorities may be informed by the sensitivity analysis.

The last step of the uncertainty analysis process is reporting. Uncertainty analysis is part of the risk assessment and should be reported in a transparent manner. It is important to list the sources of uncertainty that have been identified and document how they were identified, how each source of uncertainty has been evaluated and how they have been combined, where and how data and expert judgement have been used, what methodological approaches have been used and the rationale for choosing them, and what the results were.

It is not necessary to use all the above elements in uncertainty analysis of all risk assessments. The extent and depth of the uncertainty analysis can be scaled to the needs of the assessment and the time and resources available. In addition, the approach to each element, as well as the order in which they are conducted, may vary depending on the nature or type of each risk assessment.

14.4 Uncertainty and variability together

Most risk assessments will contain variable and uncertain inputs. In some cases it may be difficult to decide whether information relates to uncertainty and/or variability. When model parameter estimates from the scientific literature are expressed as a mean value with an associated standard deviation, it may be unclear whether this standard deviation is an expression of variability or uncertainty, or both. For example, when a growth rate is estimated from a set of growth experiments, it may not be clear whether the standard deviation in the growth rate (usually referred as a *standard error*, to denote that it refers to an estimate of a parameter) expresses uncertainty or variability. It is not sure if the growth rate is actually fixed, but cannot be determined precisely by growth experiments, or varies between the experiments but can be determined precisely. Presumably, the standard deviation expresses both. In practice, it may be important to know which characteristic is represented, and to what extent (Nauta, 2000).

When it is unclear how uncertainty and variability should be separated, there are several possible ways to proceed:

- One could test the effect of separation, assuming different 'weights' (i.e. proportional contributions) for uncertainty and variability and exploring the effect on the model outputs in several scenarios (e.g. Nauta, 2000). This will show how important it is to separate uncertainty and variability in the given situation.
- Alternatively, one might first assume that uncertainty is absent. An assumption of omniscience (pretending that everything is known) results in the remaining probability distributions necessarily describing variability. Once the variability is identified, uncertainty can then be reintroduced through scenarios by systematically varying the uncertain inputs and observing their effect on the model outputs. This approach may be quite cumbersome if there are many uncertain model inputs.
- Another way to access the potential effect of uncertainty is to identify the variable components, set their uncertain parameters to their expected value and run the model (similar to the approach described in the previous bullet point). Then the model is run as a 'mixed' model where the uncertain and variable components are simulated together using distributions as inputs. The results of the two models can then be compared to assess the potential effect of uncertainty on the model outputs and the need or otherwise to separate the two by developing a second order model.
- Cullen and Frey (1999) suggest that the relative importance of variability and uncertainty can be assessed by inspection of a two-dimensional simulation result plotted in the form of a cumulative distribution function (CDF) with confidence intervals. The mean CDF is a best estimate of variability. The confidence interval on the CDF is a best estimate of uncertainty. If the intervals are wide compared to the range of variation of the best estimate CDF, then uncertainty dominates. If the intervals are narrow, then variability dominates.
- Alternatively, Thompson and Graham (1996) provide an overview of when to select various probabilistic analysis methods depending on the policy analysis objectives.

In practice, a combination of the above approaches may be needed. For example, while uncertainty in parameter estimates can be assessed using a two-dimensional simulation model, differently structured food supply pathways may need to be assessed through different scenarios.

To illustrate the effects of variability and uncertainty consider the following situations. In the simplest case, the risk measure may be a single point probabilistic measure, e.g. the probability of at least one illness per year or the expected number of cases per year (i.e. no variability was included). This means that, if no uncertainty has been included in the risk assessment model, then the outputs are fixed values (Figure 16, top left). If uncertainty has been included in the model, then the outputs are uncertainty distributions (Figure 16a).

The risk measure may alternatively be a probability distribution capturing variability, e.g. a probability distribution of the number of adverse health events a random person might experience per year. This will be a first-order distribution if no uncertainty has been included in the model (Figure 16b), or if uncertainty and variability have been combined. If uncertainty has been included in the model and not combined with variability, then the output will be a second-order probability distribution (Figure 16c).

Thirdly, the risk measure may describe the variation in risk across a population, e.g. in different strata. That risk can, for example, be characterized as the probability of illness per serving. It can be ended up with a distribution of the variability in that probability across strata (see Section 14.1). The results can then also be stratified by graphing the variation in that probability per serving for each stratum. If the risk assessment did not include uncertainty, a single probability measure could be

used to describe the risk for each stratum (Figure 16d). If the risk assessment included uncertainty (not combined with variability), then it can be also looked at how sure about these estimates of probability per serving (Figure 16e). It is difficult to graphically compare more than two second-order distributions so, whilst it is theoretically possible to produce, for example, probability distributions of the number of illnesses per stratum over a period, if these are second-order distributions it will generally be clearer to make a comparison of an appropriate statistic (mean, 90th percentile, etc.) with attendant uncertainties.

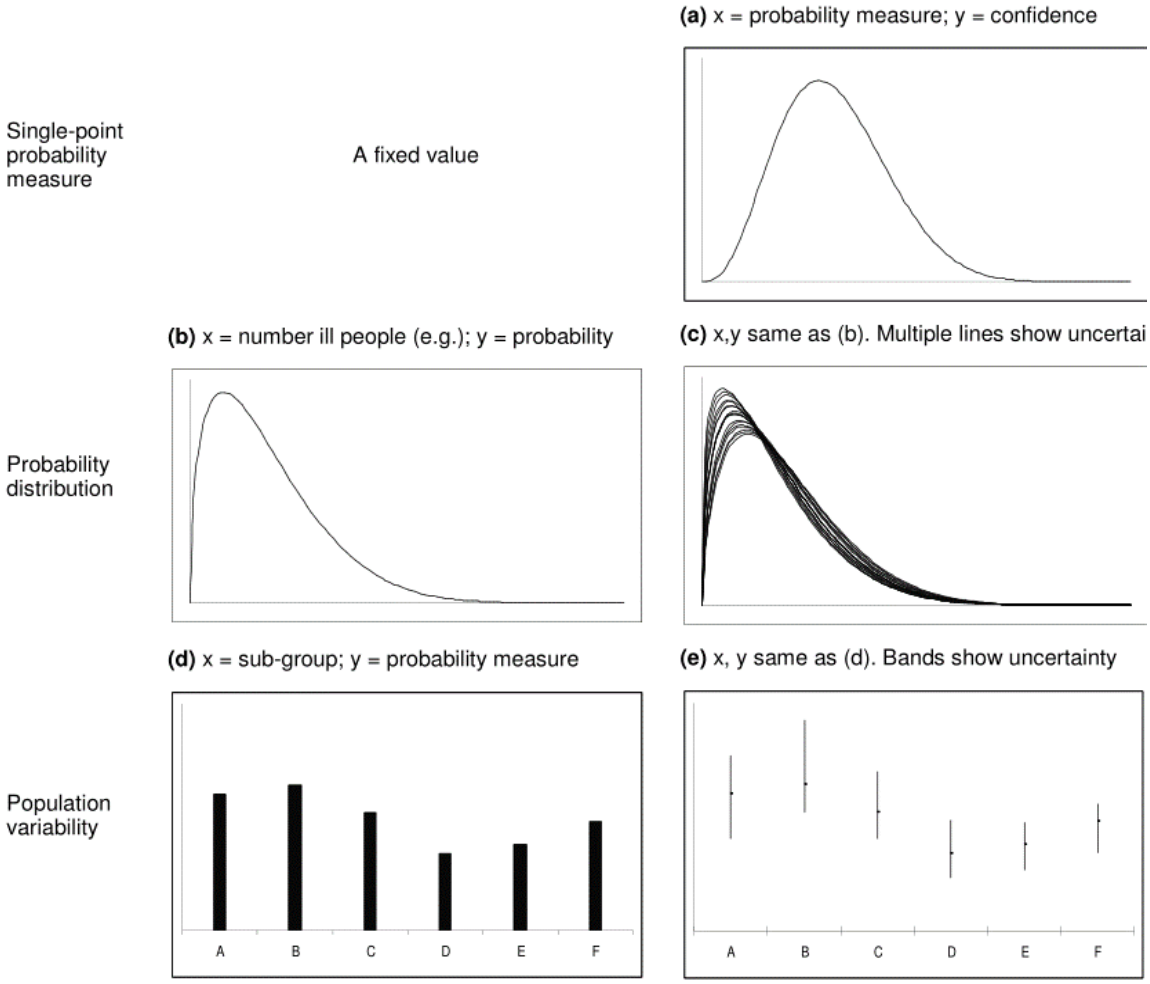

Figure 16: A matrix of various types of quantitative outputs one can produce from a risk assessment describing variability and uncertainty; variability only is shown in the graphs on the left and uncertainty and variability combined are shown in the graphs on the right.

To separate variability and uncertainty using Monte Carlo analysis, one can apply second order, or two-dimensional Monte Carlo techniques. In 'one-dimensional' simulation modelling the random realizations of the model inputs can be thought of as being arranged in a one-dimensional vector, with length equal to the number of iterations used for the model. In contrast, the two-dimensional approach can be considered as a series of such vectors, making a two-dimensional array or matrix of size ($N_v \times N_u$); the row dimension (N_v) then captures the variability in the input while the column dimension (N_u) captures the uncertainty (see Figure 2 in Pouillot *et al.*, 2007; and Figure 6 in Pouillot and Delignette-Muller, 2010). It should be noted that two-dimensional modelling is not a necessity for dealing with variability and uncertainty. In fact, "manually" investigating uncertainty and variability using for example scenario analysis can be more informative than "blindly" applying second order modelling.

15. Sensitivity analysis

Complex risk assessments may have many input and output variables that are linked by a system of equations or other model structures. Sensitivity analysis is a broad set of tools that can provide insights to risk assessors and risk managers about the relative importance of the components of a risk assessment to the risk management question (Frey, Mokhtari and Danish, 2003; Frey, Mokhtari and Zheng, 2004; Saltelli, Chan and Scott, 2008). The plausibility of important components is essential to the overall quality of the risk assessment. Changes in important components also can be expressed in terms of the influence that these inputs have on the answers to risk-management questions.

A key criterion for sensitivity analysis is that it must be relevant to a decision. Sensitivity analysis evaluates the effect of changes in model input values and assumptions on the model output, and thus on decisions that would be based on the model output. It can be used during model development to evaluate and refine model performance and can play an important role in model verification and validation. Sensitivity analysis can also be used to provide insight into the robustness of model results when making decisions.

Sensitivity analysis can also be used as an aid in identifying risk mitigation strategies or monitoring points and to focus research activities for purposes of prioritizing additional data collection or research. For these purposes, *Value of Information* (Laxminarayan and Macauley, 2012) analysis can complement sensitivity analysis methods, because the return to risk management decision-making on research and data collection expenditures depends on a variety of additional considerations, e.g. cost and time.

Microbiological risk assessment models typically have the following characteristics, which can pose substantial challenges to the application of sensitivity analysis methods:

- non-linearities;
- thresholds, e.g. below which there is no growth of a microbiological pathogen;
- discrete inputs, e.g. integer numbers of animals or ‘yes’ or ‘no’ indicators of contamination;
- incorporation of measurement error;
- variation in the scale (units and range) and shape of distributions of model inputs; and
- temporal and spatial dimensions, including dynamics, seasonality or inter-annual variability.

The relationship between model inputs and outputs should be one-to-one for effective application of sensitivity analysis methods. Ideally, a sensitivity analysis method should provide not just a rank ordering of key inputs, but also some discriminatory quantitative measure of sensitivity, such that it is possible to clearly distinguish the relative importance of different inputs (e.g. correlation). For example, are there groups of inputs among which several inputs are of comparable importance, and is there clearly a difference in importance between such groups? Statistical-based methods such as regression analysis or analysis of variance (ANOVA) produce quantitative indicators of the relative importance of different inputs, e.g. using normalized or standardized regression coefficients. Moreover, techniques such as regression analysis also provide an indication of the statistical significance of differences in sensitivity among inputs, based on confidence intervals for regression coefficients, e.g. a statistically non-significant model input implies that the model input does not have an effect on the model output. However, it should be noted that statistical tests may be able to detect very small effects, especially if the number of iterations is large, and hence any significant effect should be assessed as to its practical importance, i.e. is the effect large enough to affect risk management?

This section emphasizes sensitivity analysis in quantitative risk assessment models, although some of the techniques, e.g. exploratory methods, may apply to both quantitative and qualitative assessments.

15.1 Sensitivity analysis in qualitative risk assessment

In examining an association between a hazard and an adverse health effect, widely accepted criteria (e.g. Hill's Criteria of causation) have been established for determining whether the evidence is weak, moderate or compelling (e.g. Tomatsis, 1990). Narrative criteria may be inherently subjective, and therefore difficult to reproduce. To the extent that the criteria can be evaluated objectively, however, different assessors using the same information should be able to independently reproduce a determination of whether the criteria have been satisfied. For example, the weight of evidence for causality is stronger if detection of the association has been independently reported from multiple sources, if the strength of association is related to the level of exposure to the agent, or changes in the hazard precede changes in the observed effect. Determining whether such criteria are satisfied is evidence-based. If the results of a qualitative assessment are invariant to an accumulation of evidence regarding an association or, alternatively, to contradictory evidence, then the assessment is insensitive to the established criteria for evaluating causality. For example, in a qualitative hazard characterization, an assessment based solely on the criteria of acute health outcomes could be insensitive to information regarding known chronic sequelae. Alternatively, a qualitative hazard characterization may be highly sensitive to weak evidence regarding chronic sequelae associated with an opportunistic pathogen that rarely causes acute illness. If a qualitative risk assessment finds that a pathogen poses a negligible risk based on the assumption that the pathogen does not grow under certain environmental conditions, and new information indicates that the pathogen is capable of growing under these conditions, then the sensitivity of the findings of the risk assessment to this new information may depend on prespecified criteria, e.g. have the results been independently reproduced? Have the methods been exposed to peer review? At a minimum, the scientific basis and criteria for characterization of a qualitative risk assessment need to be sufficiently transparent to permit assessment of the effect of new information or plausible alternative assumptions on the findings.

15.2 Sensitivity analysis in quantitative risk assessment

There are several approaches to sensitivity analysis in quantitative risk assessment models. Saltelli *et al.* (2008) provide a thorough exploration of the topic, summarized below, as do Frey *et al.* (2003; 2004).

15.2.1 *Statistical methods*

Examples of statistical sensitivity analysis methods (also referred to as variance-based methods) include rank order correlations, regression analysis, ANOVA, response surface methods, Fourier amplitude sensitivity test (FAST), mutual information index (MII), and classification and regression trees (CART) (Frey, Mokhtari and Danish, 2003; Frey, Mokhtari and Zheng, 2004; Frey and Patil, 2002; Mokhtari, Frey and Jaykus, 2006). Most of these methods are applied in conjunction with, or after, a Monte Carlo simulation. Regression analysis, ANOVA, FAST and MII provide quantitative measures of the sensitivity for each input. Regression analysis requires the assumption of a model form.

15.2.2 *Graphical methods*

Graphical methods represent sensitivity typically in the form of graphs, such as scatter plots and spider plots (Eschenbach, 1992; Frey, Mokhtari and Danish, 2003). The results of other sensitivity analysis methods also may be summarized graphically, e.g. tornado charts for displaying rank order

correlation. These methods can be used as a screening method before further analysis of a model, or to represent complex dependencies between inputs and outputs. For example, such complex dependencies could include thresholds or non-linearities that might not be appropriately captured by other techniques.

15.2.3 *Evaluation of sensitivity analysis methods*

Each sensitivity analysis method provides different information (e.g. Table 5-1 in Frey, Mokhtari and Zheng, 2004) regarding sensitivities of the inputs such as the joint effect of inputs versus individual effects, small perturbations of inputs versus the effect of a range of variation, or apportionment of variance versus mutual information. Because agreement among multiple methods implies robust findings, two or more different types of sensitivity methods might be applied where practicable, to compare the results of each method and draw conclusions about the robustness of rank ordering of key inputs. Non-parametric methods, such as Spearman's rank correlation, are applicable to monotonic, non-linear models. Vose (2008) recommends the use of spider plots to illustrate the effect of individual input variables on the uncertainty of the model output.

16. Quality Assurance

The validity of any risk assessment is based on the soundness of the model structure, its inputs, the underlying assumptions and the interpretation of results. Therefore, quality assurance is a crucial element of risk assessment.

16.1 Data evaluation

Risk assessors must evaluate the quality of the data used in the analysis (see also Chapter 10), and the means of characterizing the uncertainty of any data used. The aspects listed in this section are not primarily intended for differentiating "good" from "bad" data, but rather to guide the subsequent analysis and their use in a risk assessment model.

Formalized quality control of raw data and its subsequent treatment is desirable, but also dependent on data availability and how the data are used. There is no formalized system for evaluation of data. Few generalizations can be made, but the means by which data are collected and interpreted needs to be clear. "Good" data are complete, relevant and valid: complete data are objective; relevant data are case-specific; and validation is context specific.

Data which are complete include such things as the data source and the related study information (e.g. sample size, species or strain, immune status, etc.). Characteristics of relevant data can include age of data; region or country of origin; purpose of study; analytical or data collection methods. Observations in a database should be "model free", i.e. reported without interpretation by a particular model, to allow data to be used in ways that the original investigator might not have considered. Ideally this implies that the raw data can be accessed, which may be difficult to achieve in practice. Scientific publishers are encouraging the sharing of data associated with publications, where possible, and independent data repositories have also been created; see for example http://foodrisk.org or https://www.combase.cc.

Valid data are those that agree with others in terms of comparable methods and test development. In general, for dose-response modelling, for example, human data need less extrapolation and are preferred to animal data, which in turn are preferable to in vitro data. Data on the pathogen of concern are preferred to data on surrogate organisms, which should only be used when proven to be valid (NACMCF, 2010).

The current recommended practice is to consider any available data as potentially useful. Whether data should be eliminated from the risk assessment depends on the purpose and stage of the assessment. Small data sets or those with qualitative values may be useful in the early stages of a risk assessment. The later stages of risk assessment may include only those data that meet a particular quality standard. Excluding data from the analysis should be based on predefined criteria, e.g. age of the data set, geographic representativeness, etc., and not based solely on statistical criteria (e.g. Section 16.1.2). If the data are extremely heterogenous or contain outliers, they may be stratified according to suitable criteria. This practice should provide increased insight rather than information loss.

Sources of data may come from the peer-reviewed or non-peer-reviewed literature. Although peer reviewed data are generally preferable, they also have some important drawbacks (see also Section 10.1). Access to the peer reviewed literature may be restricted especially for developing countries, although open-access publications and Research4Life (see Section 3.5.2), for example, are helping to address some of these limitations. Peer reviewed data may be missing important methodological details (e.g. sample preparation and characteristics), are usually presented in an aggregated form, and may not provide the level of detail necessary for uncertainty analysis. Quality control of the

measurement process may be poorly documented. The potential for publication bias should not be ignored, as 'replication studies' may not provide enough novelty for publishers and hence may only get published through conference presentations, reports or other formats. The analyst might wish to add information from other sources for any of these reasons. The quality of any data used should be explicitly reviewed, preferably by independent experts, and any concerns regarding data quality should be explicitly noted.

The results of any risk assessment are conditional on the data and information used to develop the risk model. Any risk assessment should summarize the primary strengths and limitations of the data, methods, and analyses used. Typically, these analyses require risk assessors to synthesize and draw inferences from disparate data sources not specifically or originally intended for use in risk assessment. In some cases, this requires the use of unconventional or non-routine methods that might be highlighted for close scrutiny, to ensure that they are reasonable and correctly applied.

16.1.1 *Data collection*

Suitable data for microbiological risk assessment may be sparse. Assessors should initially collect all reasonably obtainable data consistent with the assessment objective, and subsequently investigate the quality of different data sources. When collecting data for input distributions, several issues should be considered to evaluate data quality. The following considerations apply to any data, including information elicited from experts.

Risk assessors should ideally have access to raw, un-summarized data (e.g. EcoSure, 2008). Statistical methods such as quantile-quantile plots and skewness-kurtosis plots (Cullen and Frey, 1999) can be useful to identify suitable parametric distributions, if the raw data contain sufficient observations. Alternatively, empirical distributions or non-parametric simulation methods can be used to characterize input distributions. Graphical methods are generally preferred over statistical tests (e.g. Goodness-of-Fit) which are affected by the size of the data set. Large sample sizes can identify statistically significant deviations from the hypothesised distribution, even though these deviations may be of little practical importance.

Raw data are frequently inaccessible and results are often reported as aggregated summary statistics (e.g. estimated mean, standard deviation or standard error). It may be necessary to obtain information on the assumed distribution of the underlying data, together with the sample size to develop a distribution from data summary statistics.

It is useful to collect as much background information on the data sources as possible, such as the year of completion, country of origin, the type of sample, possible transformation of the data, methods of analysis, microbiological strain and population demographics. This information could be important about treatment or use of the data or to support the decision on whether to include these data in the model. An example is given below.

This example is from a Danish risk assessment for *Campylobacter jejuni* in chicken (Christensen *et al.*, 2001). Quantitative data were needed to describe the relative change in pathogen concentration over a given step in a poultry slaughterhouse. Data from foreign studies were applied to assess the efficacy of the wash and chiller process in reducing the pathogen levels on chicken carcasses because Danish data were unavailable. Data for the microorganism of interest were available, but the data were obtained from different sample units (neck skin samples, whole carcass wash, and swab samples). This mix of sample types all reflected surface contamination of chicken carcasses. The risk assessors assumed that the relative reduction in pathogen concentration over the process was independent of the type of surface measure. The slopes of the lines shown in Figure 17 reflect

differences in log-concentration over the process. Since all the slopes appear to be similar (though not identical), all data sets were used in describing the reduction over the 'wash + chiller' process.

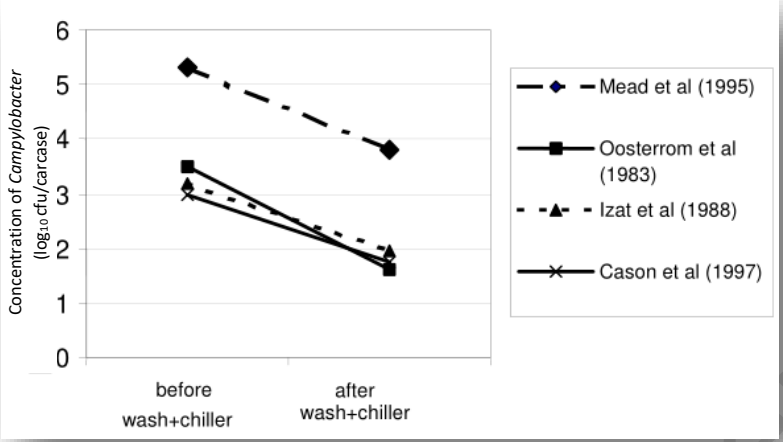

Figure 17: The influence of a selected slaughterhouse process on the Campylobacter concentration on chicken carcasses. The change in pathogen concentrations (expressed as log CFU per carcase) before and after the process is represented by a line connecting data points originating from the same study.

Data for the specific microorganism under study may not always be available or of suitable quantity and quality. Data from a surrogate microorganism may be used, provided that the surrogate behaves similarly under the process of interest, e.g. generic *E. coli* to estimate cross-contamination during slaughter procedures. Data from different surrogate organisms could be used to model different steps in the same model, based on data availability and suitability. Sampled data with different units, e.g. absolute concentration or change in concentration, can be used to describe the same process, as the example above illustrates. Depending on how the data are used in the model, e.g. describing a change in concentration over a step or describing the concentration level, different parameters may be evaluated in a sensitivity analysis to ensure data quality objectives are satisfied.

In some cases, the available data may not be representative of the population of interest. These data may be excluded from the analysis or incorporated with appropriate adjustment. The bases for decisions regarding the treatment of non-representative data are context specific and need to be clearly articulated. For example, data from a particular source may be considered non-representative for the purposes of providing an estimate of central tendency (e.g. the mean) but may nevertheless be useful for the purposes of characterizing the spread of an input distribution (e.g. plus or minus an order of magnitude).

16.1.2 *Sorting and selecting data sources*

After collecting potentially suitable data sets, the risk assessor should evaluate each critically and select the data that will provide the most appropriate model input for the specific purpose (e.g. contamination level, contamination prevalence or changes during processing). Plotting the parameter of interest with the 95% confidence intervals provides a useful overview (see Figure 18).

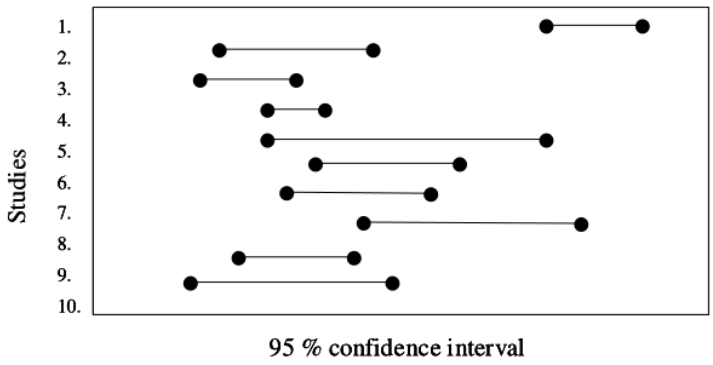

Figure 18: Example of an overview of data from different studies, with model input parameter 95% confidence intervals.

Both subjective and statistical criteria may be applied in selecting the suitable data sets for incorporation into the risk assessment. Subjective evaluation criteria may include the representativeness of the geographical and temporal properties of the study. If study 1 in Figure 18 is the only study conducted outside the country of interest, and it is significantly different from the rest (based on statistical criteria), this data set could be excluded. If the 10 studies all originate from the same country, but are reported by different laboratories, the differences may be due to variability between the laboratories or specific sampling context and the assessor might decide to incorporate all studies in the model. Irrespective of the decision taken, the rationale should be documented.

16.2 Model Quality Assurance

Models should be both verified and validated and may also be anchored (calibrated). Model verification is achieved by auditing the model to ensure that it operates as intended by the developer. Anchoring and calibration are techniques to adjust the model to approximate observed data. Model validation can be defined as demonstrating the accuracy of the model for a specified use. Model verification should precede model validation. If the model is to be both anchored and validated, using a withheld test portion of the independent data, then anchoring should precede model validation.

16.2.1 *Model verification*

Verification includes checking of the software code used to implement the model. Verification requires that the model be suitably documented. All data, methods, assumptions and tools used should be clearly described, so that the model can be independently reproduced. A well-organized model structure facilitates verification.

The following questions may be useful for those seeking to verify a model:

- Are the analytical equations correctly derived and free of error? If approximations are used, then under what assumptions do they hold and are those assumptions always met?
- Is the computerized version of the analytical model correctly implemented? What, if any, are the limits of the implementation?
- Are the inputs correctly specified?
- Do the units of measurement (e.g. CFU or log CFU) propagate correctly through the model?
- Is the model internally consistent? For example, if an assumption is made in one part of the model, is it consistently applied throughout the model? Is there consistency within the model between the intermediate outputs and inputs?

- Are errors in any computational step flagged appropriately, or could they result in inappropriate values being propagated through the model?
- Are the intermediate outcomes and end results evaluated to be realistic?

It may be difficult in some cases to do a line-by-line verification of computer code, especially for large models. The verification of any computer code will be facilitated if good software engineering practices (e.g. Pressman, 2005) are followed, including clear specification of databases, development of a software structure design prior to coding, version control, modular design, clear specification of interfaces between components of a model, and good communication among project teams when different individuals are developing different components of a model. Literate programming techniques (Knuth, 1992) can also be useful for this purpose as they allow embedding of the model code in the documentation; a range of tools for various programming languages and environments are available.[22] Model documentation and peer review are critical aspects of the verification process.

16.2.2 *Model anchoring or calibration*

Anchoring is a technique in which the model is adjusted, or calibrated, to be more compatible with observed data. For example, model parameters may be adjusted to achieve agreement between model predictions and observed data, e.g. predicted versus actual number of illnesses per year attributed to the hazard and the food. As noted above, if the model is to be both anchored and validated, using a withheld test portion of the independent data, then anchoring should precede model validation.

Anchoring is a generally accepted practice in health risk assessment and environmental modelling, and has been employed in one fashion or another in various risk assessments (CFSAN and FSIS, 2003; FAO/WHO, 2005; FSIS, 2001, 2005). Data from outbreaks could be considered as the ultimate 'anchor' for dose-response models and are also an important way to validate risk assessments. This is because the dose ingested by different consumers involved in an outbreak is likely to be more similar than the doses associated with sporadic cases. Since anchoring requires some data, it may compromise efforts to validate the model in situations without sufficient data to support both activities. A common approach in statistics and machine learning (e.g. neural networks, etc.) is to separate a data set into two independent components: training and test data. The training data are used to fit the model and estimate the model parameters, while the test data are used to independently check the predictions of the model against previously unseen observations. In general, anchoring approaches that weigh model inputs in proportion to their likelihood in light of the observed data are superior to using simple adjustment factors or censoring input values that are incompatible with the observed data (NAS, 2002; Williams, Ebel and Vose, 2011b). Whatever the anchoring approach, considerable care must be taken to ensure that the adjustment procedure is well reasoned and transparent.

16.2.3 *Model validation*

Model validation can be defined as demonstrating the accuracy of the model for a specified use. Accuracy is the absence of systematic and random error, commonly known as trueness and precision, respectively. Models are always incomplete representations of the system they are intended to model, but they can still be useful. General information on working with mathematical models can be found in various theoretical and applied textbooks. Doucet and Sloep (1992) give a good introduction to model testing. These authors discriminate between models shown to be worthy of our belief (i.e. plausible) and models shown to be true. McCullagh and Nelder's book on generalized linear models (1989) is a valuable resource on statistical modelling methods, and

[22] https://en.wikipedia.org/wiki/Literate_programming accessed 5 December 2018

describes some general principles of applying mathematical models, underlining three key principles: (i) All models are wrong, but some are more useful than others; (ii) Do not fall in love with one model to the exclusion of others; and (iii) Thoroughly check the fit of a model to the data.

Law (2014), in addressing the issue of building valid, credible and appropriately detailed simulation models, considers techniques for increasing model validity and credibility. Model validation procedures should be aimed at answering questions like: (i) Does the model make sense; (ii) Does the model respond in an appropriate manner to changes in input assumptions; and (iii) Do predictions respond in an appropriate manner to changes in the structure of the analysis? These processes are also referred to by some as a 'reality check', 'laugh test' or 'confidence building'.

Model validation is highly dependent on the risk-management question, and the degree of validation required should be proportionate to the stakes of the decision. Model validation involves demonstrating the accuracy of the model for a specified use and there are different aspects to model validation. Dee (1994, 1995) identified four major aspects associated with model validation: (i) Conceptual validation; (ii) Validation of algorithms; (iii) Validation of software code; and (iv) Functional validation. These are described below.

Risk assessment, like any type of problem solving is cyclical in nature. Defining the problem, considering alternative solutions, and implementing a solution all lead to the need to assess the effectiveness of the chosen solution. The cycle may repeat based on that assessment. No risk assessor should think their job is done after a risk management decision is made. The risk assessor may begin planning how they will assess the validity of the predictions of their model in the context of the risk management option selected. This assessment of validity may not occur until years after risk management options are implemented.

Conceptual validation concerns the question of whether the model accurately represents the system under study. Was the simplification of the underlying biological process in model steps realistic, i.e. were the model assumptions credible? Usually, conceptual validation is largely qualitative and is best tested against the opinion of experts with different scientific backgrounds. Different models with various conceptual bases can be tested against each other within a Bayesian framework, using Bayes factors, or some other information criterion (Kass and Raftery, 1995). Experimental or observational data in support of the principles and assumptions should be presented and discussed. With respect to dose-response modelling, the concepts described in Section 6.3 are a minimum set of assumptions representing the consensus opinion of a broad group of experts who contributed to the original FAO guidelines on hazard characterization (FAO/WHO, 2003). These are based on mechanistic reasoning and are supported by some experimental evidence. As such, they are considered to be currently the best basis for dose-response modelling studies.

Algorithm validation concerns the translation of model concepts into mathematical formulae. It addresses questions such as: Do the model equations represent the conceptual model? Under which conditions can simplifying assumptions be justified? What effect does the choice of numerical methods for model solving have on the results? and: Is there agreement among the results from use of different methods to solve the model?

Software code validation concerns the implementation of the model in a computer language. Good programming practice (i.e. modular and fully documented) is an essential prerequisite. Specific points for attention are the possible effects of machine precision and software-specific factors on the model output. For this reason, open-source software and models implemented in a computing language (e.g. R, Python, C++, etc.) may be preferable to those implemented in a proprietary

software program, as all computational steps can be inspected if needed. Internal error reports of the software are important sources of information, as well as evaluation of intermediate output.

Functional validation concerns checking the model against independently obtained observations. Ideally, it is evaluated by obtaining pertinent real-world data, and performing a statistical comparison of simulated outcomes and observations (Ebel and Williams, 2019). This requires more detailed information than is usually available, especially if data are also used for anchoring (Section 16.2.2). It may be possible to compare results from risk assessment studies with independently obtained epidemiological estimates of disease incidence. Such data cannot validate a dose-response model *per se* but may produce valuable insights, especially if the predictions do not closely match epidemiological observations. Most studies to date have considered that a range check of estimated risks and observed incidences was sufficient "validation" of the model.

Credibility of results can also be established by demonstrating that different sources of data are consistent with output values. These might include intermediate outputs. Cassin *et al.* (1998) provide a good example of such comparisons. When making such comparisons, the different nature of vehicle, microbiological hazard and processes must be accounted for. It should be noted that if the model output does not agree with the observations, it might not necessarily be that the model is wrong. It may be that the observation itself was influenced by an unknown factor (e.g. microbiological methodological insensitivity) or the underestimation of foodborne illness associated with current epidemiological data. There may also be a variety of different influences acting in concert to cause the differences in the results.

Close agreement between an initial risk-modelling effort and independent validation data would be fortuitous. Agreement between the model output and validation data may be coincidental, however, and would not necessarily indicate that all of the intermediate model components are accurate. Typically, model development and refinement are iterative. Whether model validation or anchoring is considered, the credibility of the model may be strengthened by having multiple points at which the model can be compared to observed data. In general, the scientific credibility of a model is strengthened if consistent results are derived from different relevant data sources (e.g. laboratories, regions) or types (observational or experimental), or a combination. The required degree of relevance and consistency is a context-specific judgement. The tolerance for inconsistent answers depends on what constitutes an 'important' difference with respect to changes in model results. In the risk assessment context, an important difference in model results is one that would significantly modify the risk management decision under the relevant decisional criteria.

There are situations in which it may be difficult, or practically impossible, to completely validate a model. For example, because risk assessment models are often attempting to predict low probability events, it can be difficult to obtain an independent data set of sufficient sample size to make a sensible comparison of predictions versus observations. It may be possible to validate components of the model even in such situations. For example, it may be possible to validate portions of the model that deal with a particular exposure pathway by making measurements of contaminant levels in specific foods.

In many cases, there may be insufficient or no independent data with which to compare model predictions. In these situations, alternatives to validation include: (i) Screening procedures to identify the most important model inputs and pathways; (ii) Sensitivity analysis to identify the most important inputs or groups of inputs; (iii) Uncertainty analysis to evaluate the effect of uncertainty in model inputs with respect to predictions; (iv) Comparison among predictions of different models;

and (v) Evaluation of sensitivity of results to different assumptions regarding scenarios, model boundaries, model resolution and level of detail.

While none of these techniques provides a direct validation of the model, each of these techniques provides insight into the sensitivity of the model predictions to key assumptions regarding the analysis. The response of the predictions to these procedures can be evaluated with respect to prior expectations, comparison with analogous systems, and theoretical justifications.

16.3 Comparison with epidemiological data

To make a valid comparison with a foodborne pathogen risk estimate, at least three factors need to be considered when deriving an epidemiological estimate from human surveillance data (Powell, Ebel and Schlosser, 2001). These factors are (i) Cluster-weighted rate of illness; (ii) Adjustment of surveillance data to account for under-reporting; and (iii) Etiological fraction attributable to food products. These three factors are discussed in more detail below.

If the risk assessment estimates the incidence of illness at the national level, the epidemiological estimate will need to extrapolate the rate of illness beyond the surveillance area to permit comparison at the national level. In this case, the raw reported rate in each surveillance area may be weighted by the population of the region represented by the area (e.g. state population size) to obtain a weighted average rate of illness (e.g. cases per 100,000 in the national population). If multiple years of surveillance data are available, then the data can be used to characterize year-to-year variability in the rate of illness.

Estimating the actual incidence of illness requires adjustment for recognized sources of underreporting in human surveillance data (Scallan *et al.*, 2011; Williams, Ebel and Vose, 2011b). For example, some ill persons do not seek medical care, physicians do not obtain stool specimens from all patients, laboratories do not culture all stool samples for the pathogen of concern, and some proportion of the lab results are false negatives. If estimates are available on the proportion of cases at each step in the reporting process, the negative binomial distribution can be used in sequential fashion to estimate the number of cases missed at each step. In some cases, the proportions may be dependent on the nature or severity of symptoms. For example, a person with bloody diarrhoea may be more likely to seek medical care than one with non-bloody diarrhoea. In this case, the proportion of cases with different levels of symptoms must be estimated prior to accounting for the number of cases missed at each step, and the adjusted symptom-specific estimates are summed to estimate the total number of cases (Hall *et al.*, 2008). In general, the degree of under-reporting tends to be substantial (WHO, 2015), and varies among countries and between regions within countries (Scallan *et al.*, 2011).

The etiological fraction refers to the proportion of cases attributable to an exposure pathway or a specific food product (Greig and Ravel, 2009; Mullner *et al.*, 2009; Painter *et al.*, 2013; Pires, 2013; Pires *et al.*, 2009). If the scope of the risk assessment is limited to a particular food product, then the proportion of cases due to other exposure pathways (e.g. other foods, drinking water) needs to be subtracted from the overall estimate of illness obtained from the human surveillance data. In general, empirical data on the etiological fraction are scarce. It may be possible, however, to specify a range of uncertainty on the basis of expert judgement (e.g. Vally *et al.*, 2014).

16.4 Extrapolation and robustness

Model robustness refers to the performance of the model when its assumptions are violated. In this context, assumptions include model form and model inputs. Extrapolating model results to other settings may involve many forms of extrapolation (e.g. from the present to the future, from one

geographical region to another, from one microorganism to another, from animals to humans, from clinical trial subjects to the general population, from one population to another, from available data to values beyond the observed range , from experimental settings to operational environments). Some extrapolations can be made with relative confidence, while others can not. Some degree of extrapolation may be inevitable, since the demands of risk management may outstrip the supply of relevant science. The importance of various forms of extrapolation made in risk assessment needs to be considered and, to the extent feasible and relevant to the decision at hand, characterized in a clear manner, either quantitatively or qualitatively.

Extrapolation is explicit when the selected values of model inputs are outside the range of values used to calibrate or validate the model, or both. However, there can also be hidden extrapolation. A hidden extrapolation occurs for a combination of values of two or more model inputs such that these values individually are enclosed by ranges used for calibration and validation, but for which the specific combination was not included or approximated during calibration or validation. Thus, simple range checks on each input will not guarantee that a hidden extrapolation cannot occur. Hidden extrapolation would typically be more of a problem for a system in which there are highly sensitive interactions among inputs or when model inputs are highly correlated.

A model that is calibrated to a narrow range of values for each input may not be robust when applied to sensitivity or uncertainty analysis. The use of ranges or distributions, rather than point estimates, could lead to hidden or explicit extrapolations of the model. Situations may also arise in which some iteration of Monte Carlo simulation give division by zero or unbounded result errors. Such problems can often be solved by investigating model assumptions, checking model inputs, or adding error trapping in the software. Problems such as these can arise in practice, particularly when working with a model or computer code that someone else developed and for which documentation may be inadequate.

A model is considered to be robust if it responds in a reasonable manner to variation in input values, while at the same time not being easily subject to singularity points or other structural issues that lead to substantial magnification of errors in input values, whether because of uncertainty or user error. A model that is based on sound theory might be used with more confidence compared with a purely empirical model that is essentially “curve fitting”.

16.5 Credibility of the risk assessment

Documentation, validation, and review are necessary criteria for the credibility of a risk assessment. None of these criteria is sufficient by itself, however, as credibility depends on all three criteria being satisfied in a manner that is proportionate to the stakes of the decision. Documentation and scientific review are discussed below and validation has already been discussed in Section 16.2.3.

16.5.1 *Risk assessment documentation*

Risk assessment documentation should serve both technical and non-technical readers. One way to address this need is to provide a technical document with all modelling details and a less technical interpretive summary.

Risk assessment documentation must enable the analysis to be independently reproduced. Modern programming tools, free and open-source software, and sharing of risk assessment model code may assist in this aim. The principle of transparency also requires that the source or basis for model inputs and assumptions be clearly stated, e.g. by references to scientific literature, evaluation criteria or expert judgement. The expectation for risk assessment documentation should be reasonable, however, because in some cases, assumptions may be based on common knowledge or

generally accepted practices in the field. For example, the log-normal distribution is commonly assumed for modelling variables that are the product of several other variables. Because risk assessments are difficult to fully validate, and because such assessments are used to inform public health decision-making at various levels, including local, national, and international, it is critically important that the information used for the assessment, including the model, be accessible for review by experts and the lay public (e.g. FAO/WHO, 2009c, 2009d).

The information in the documentation of a risk assessment should include:

1. Data or references to data sources;
2. Scenarios, including the temporal and spatial aspects of the exposure scenarios, the specific hazards addressed, the specified pathogens included, exposed populations and exposure pathways;
3. The analytical model used for analysis, including the theoretical or empirical basis;
4. Discussion and comparison of alternative model formulations, and justification for choices made regarding model structure;
5. Assumptions regarding values assigned to model inputs, including point-estimates, ranges and distributions;
6. Model verification, including assessment of results from sensitivity and uncertainty analysis;
7. Model anchoring (calibration);
8. Model validation; and
9. Computer implementation of the analytical model, including software design.
10. An interpretive summary that is understandable by the risk manager.

16.5.2 *Scientific peer review*

The credibility of risk assessment results can be improved by the process used to develop the results. Peer and public review of risk assessment results are an essential part of the process, but each type of review generates distinct and sometimes conflicting demands that should be addressed on their own terms.

Morgan and Henrion (1992) identify exposure to peer review as a basic tenet of good policy analysis. The focus of a scientific peer review is highly dependent on the risk management question that the risk assessment is intended to inform. Without reference to a well-defined and specific risk management question, peer review of a risk assessment may fail to focus on the particular uncertainties that are most likely to influence the risk management decision. For example, if the risk management question is "What is the likelihood that a specific pathogen occurs in a particular food production process?" then data gaps and other uncertainties regarding post-production processes are irrelevant to the decision. Peer review comments regarding the scope of the risk assessment, while potentially useful for future risk assessments, are not relevant to the adequacy of the risk assessment under review to inform the risk management decision for which it was intended. If a risk assessment has multiple objectives, peer review may help to identify which objectives an assessment satisfies, since an assessment that is adequate to inform one decision may be insufficient to support another. A thorough review can be difficult and time consuming for a complex risk assessment, even if the documentation is adequate. In the case of large, complex risk assessments, a thorough review may require a multidisciplinary team and a significant budget, e.g. the NAS review (NAS, 2002) of the FSIS risk assessment of *E. coli* O157 in ground beef (FSIS, 2001). The substantive and procedural benefits of peer review should therefore be balanced by time and resource considerations. The level and extent of review should be proportionate to the stakes of the decision, taking into consideration the need for immediate action in the event of actual public health emergencies.

Bibliography

den Aantrekker, E.D., Beumer, R.R., van Gerwen, S.J.C., Zwietering, M.H., van Schothorst, M. & Boom, R.M. 2003. Estimating the probability of recontamination via the air using Monte Carlo simulations. *International Journal of Food Microbiology*, 87(1–2): 1–15.

Abramowitz, M. & Stegun, I.A. 1972. *Handbook of mathematical functions*. 9th edition. New York, USA, Dover.

ACMSF. 2012. ACMSF Approaches to Microbiological Risk Assessment: Discussion Paper - Risk Assessment Outputs. No. ACM/1065. Advisory Committee on the Microbiological Specifications of Food. (also available at https://acmsf.food.gov.uk/sites/default/files/mnt/drupal_data/sources/files/multimedia/pdfs/committee/acm_1065.pdf).

Alber, S.A. & Schaffner, D.W. 1992. Evaluation of data transformations used with the square root and schoolfield models for predicting bacterial growth rate. *Applied and Environmental Microbiology*, 58(10): 3337–3342.

Al-Zeyara, S.A., Jarvis, B. & Mackey, B.M. 2011. The inhibitory effect of natural microflora of food on growth of *Listeria monocytogenes* in enrichment broths. *International Journal of Food Microbiology*, 145(1): 98–105. https://doi.org/10.1016/j.ijfoodmicro.2010.11.036

Anderson, R.M. & May, R.M. 1992. *Infectious Diseases of Humans: Dynamics and Control*. OUP Oxford. 772 pp.

WHO. 2018. Whole genome sequencing for foodborne disease surveillance: landscape paper. Geneva, Switzerland, WHO. (also available at http://www.who.int/foodsafety/publications/foodborne_disease/wgs_landscape/en/).

Arrow, K.J. 1963. Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review*, 53(5): 941–973.

Arthur, T.M., Bosilevac, J.M., Brichta-Harhay, D.M., Guerini, M.N., Kalchayanand, N., Shackelford, S.D., Wheeler, T.L. & Koohmaraie, M. 2007. Transportation and Lairage Environment Effects on Prevalence, Numbers, and Diversity of *Escherichia coli* O157:H7 on Hides and Carcasses of Beef Cattle at Processing. *Journal of Food Protection*, 70(2): 280–286. https://doi.org/10.4315/0362-028X-70.2.280

Bailey, N.T.J. 1975. *The mathematical theory of infectious diseases and its applications.*. Second edition. New York, Hafner Press.

Baranyi, J., Buss da Silva, N. & Ellouze, M. 2017. Rethinking Tertiary Models: Relationships between Growth Parameters of Bacillus cereus Strains. *Frontiers in Microbiology*, 8. https://doi.org/10.3389/fmicb.2017.01890

Baranyi, J., Ross, T., McMeekin, T.A. & Roberts, T.A. 1996. Effects of parameterization on the performance of empirical models used in `predictive microbiology'. *Food Microbiology*, 13(1): 83–91. https://doi.org/10.1006/fmic.1996.0011

Bassett, J., Jackson, T., Jewell, K., Jongenburger, I. & Zwietering, M.H. 2010. Impact of Microbial Distributions on Food Safety. International Life Sciences Institute (ILSI) Europe. (also available at http://ilsi.org/publication/impact-of-microbial-distributions-on-food-safety/).

Bassett, J., Nauta, M.J., Lindqvist, R. & Zwietering, M.H. 2012. Tools for Microbiological Risk Assessment. International Life Sciences Institute Europe. (also available at http://ilsi.org/publication/tools-for-microbiological-risk-assessment/).

Batz, M., Hoffmann, S. & Morris, J.G. 2014. Disease-outcome trees, EQ-5D Scores, and estimated annual losses of Quality-Adjusted Life Years (QALYs) for 14 foodborne pathogens in the United States. *Foodborne Pathogens and Disease*, 11(5): 395–402. https://doi.org/10.1089/fpd.2013.1658

Berends, B.R., Van Knapen, F., Mossel, D.A.A., Burt, S.A. & Snijders, J.M.A. 1998. Impact on human health of *Salmonella* spp. on pork in The Netherlands and the anticipated effects of some currently proposed control strategies. *International Journal of Food Microbiology*, 44(3): 219–229. https://doi.org/10.1016/S0168-1605(98)00121-4

Bergold, J. & Thomas, S. 2012. Participatory Research Methods: A Methodological Approach in Motion. *Historical Social Research / Historische Sozialforschung*, 37(4 (142)): 191–222.
Berjia, F.L., Poulsen, M. & Nauta, M.J. 2014. Burden of diseases estimates associated to different red meat cooking practices. *Food and Chemical Toxicology*, 66: 237–244. https://doi.org/10.1016/j.fct.2014.01.045
Biosecurity Australia. 2016. *Biosecurity Import Risk Analysis Guidelines 2016: Managing biosecurity risks for imports into Australia*. Canberra, ACT, Australia, Department of Agriculture and Water Resources. (also available at http://www.agriculture.gov.au/SiteCollectionDocuments/bira-guidelines-2016.pdf).
Bogaardt, M.J., Mangen, M.-J.J., de Wit, G.A., Nauta, M.J. & Havelaar, A.H. 2004. Controlling *Campylobacter* in the chicken meat chain - Towards a decision support model [Beheersing van campylobacter in de ippenvleesketen-naar een eslissingsondersteunend model]. , p. 56. No. 250911005. Bilthoven, The Netherlands, RIVM [Rijksinstituut voor Volksgezondheid en Milieu]. (also available at https://www.rivm.nl/bibliotheek/rapporten/250911005.html).
Bogen, K.T. 1994. A Note on Compounded Conservatism. *Risk Analysis*, 14(4): 379–381. https://doi.org/10.1111/j.1539-6924.1994.tb00255.x
de Boor, C. 2001. *A Practical Guide to Splines*. New York, Springer. 348 pp.
Brul, S., van Gerwen, S.J.C. & Zwietering, M.H., eds. 2007. *Modelling microorganisms in food*. Woodhead publishing in food science, technology and nutrition. Abington Cambridge, Woodhead Publ. Limited. 294 pp.
Buchanan, R.L. 1993. Predictive food microbiology. *Trends in Food Science & Technology*, 4(1): 6–11. https://doi.org/10.1016/S0924-2244(05)80004-4
Buchanan, R.L., Damert, W.G., Whiting, R.C. & van Schothorst, M. 1997. Use of Epidemiologic and Food Survey Data To Estimate a Purposefully Conservative Dose-Response Relationship for *Listeria monocytogenes* Levels and Incidence of Listeriosis. *Journal of Food Protection*, 60(8): 918–922. https://doi.org/10.4315/0362-028X-60.8.918
Buchanan, R.L., Gorris, L.G.M., Hayman, M.M., Jackson, T.C. & Whiting, R.C. 2017. A review of Listeria monocytogenes : An update on outbreaks, virulence, dose-response, ecology, and risk assessments. *Food Control*, 75: 1–13. https://doi.org/10.1016/j.foodcont.2016.12.016
Buchanan, R.L., Smith, J.L. & Long, W. 2000. Microbial risk assessment: dose-response relations and risk characterization. *International Journal of Food Microbiology*, 58(3): 159–172. https://doi.org/10.1016/S0168-1605(00)00270-1
Burgman, M.A. 2015. *Trusting Judgements: How to Get the Best out of Experts*. Cambridge University Press. 215 pp.
Burmaster, D.E. & Harris, R.H. 1993. The Magnitude of Compounding Conservatisms in Superfund Risk Assessments. *Risk Analysis*, 13(2): 131–134. https://doi.org/10.1111/j.1539-6924.1993.tb01058.x
Busschaert, P., Geeraerd, A.H., Uyttendaele, M. & Van Impe, J.F. 2011. Hierarchical Bayesian analysis of censored microbiological contamination data for use in risk assessment and mitigation. *Food Microbiology*, 28(4): 712–719.
CAC. 1999. *CAC/GL 30-1999: Principles and Guidelines for the Conduct of Microbiological Risk Assessment (Amendments 2012, 2014)*. Codex Alimentarius Commission. (also available at http://www.fao.org/fao-who-codexalimentarius/home/en/).
CAC. 2000. Report of the 33rd Session of the Codex Committee on Food Hygiene (CCFH). Washington, DC, USA, Codex Alimentarius Commission. (also available at http://www.fao.org/fao-who-codexalimentarius/meetings/archives/en/).
CAC. 2007. *CAC/GL 62-2007: Working Principles for Risk Analysis for Food Safety for Application by Governments*. Codex Alimentarius Commission. (also available at http://www.fao.org/fao-who-codexalimentarius/home/en/).

CAC. 2008. *CAC/GL 63-2007: Principles and Guidelines for the Conduct of Microbiological Risk Management (MRM)*. Codex Alimentarius Commission. (also available at http://www.fao.org/fao-who-codexalimentarius/home/en/).
CAC. 2018. *Codex Alimentarius Commission Procedural Manual*. 26th edition. Codex Alimantarius Commission. (also available at http://www.fao.org/fao-who-codexalimentarius/home/en/).
CAC. 2019. *Codex Alimentarius Commission Procedural Manual*. 27th edition. Codex Alimantarius Commission. (also available at http://www.fao.org/fao-who-codexalimentarius/home/en/).
Cadavez, V.A.P., Campagnollo, F.B., Silva, R.A., Duffner, C.M., Schaffner, D.W., Sant'Ana, A.S. & Gonzales-Barron, U. 2019. A comparison of dynamic tertiary and competition models for describing the fate of Listeria monocytogenes in Minas fresh cheese during refrigerated storage. *Food Microbiology*, 79: 48–60. https://doi.org/10.1016/j.fm.2018.11.004
Cardoen, S., Van Huffel, X., Berkvens, D., Quoilin, S., Ducoffre, G., Saegerman, C., Speybroeck, N., Imberechts, H., Herman, L., Ducatelle, R. & Dierick, K. 2009. Evidence-Based Semiquantitative Methodology for Prioritization of Foodborne Zoonoses. *Foodborne Pathogens and Disease*, 6(9): 1083–1096. https://doi.org/10.1089/fpd.2009.0291
Cassin, M.H., Lammerding, A.M., Todd, E.C.D., Ross, W. & McColl, R.S. 1998. Quantitative risk assessment for *Escherichia coli* O157:H7 in ground beef hamburgers. *International Journal of Food Microbiology*, 41: 21–44.
Cassin, M.H., Paoli, G.M. & Lammerding, A.M. 1998. Simulation Modeling for Microbial Risk Assessment. *Journal of Food Protection*, 61(11): 1560–1566. https://doi.org/10.4315/0362-028X-61.11.1560
Cassin, M.H., Paoli, G.M., McColl, R.S. & Lammerding, A.M. 1996. A Comment on 'Hazard Assessment of *Listeria monocytogenes* in the Processing of Bovine Milk'. *Journal of Food Protection*, 59(4): 341–343.
CCFFP. 2011. Agenda Item 3 (Comment) - Thailand information paper: On estimating the risk of developing histamine poisoning from the consumption Thai fish sauces. No. FFP/31 CRD 18. Codex Alimentarius Commission. (also available at http://www.fao.org/tempref/codex/Meetings/CCFFP/ccffp31/CRD/CRD_18_Thailand.pdf).
CFSAN & FSIS. 2003. Quantitative Assessment of Relative Risk to Public Health from Foodborne Listeria monocytogenes Among Selected Categories of Ready-to-Eat Foods. Center for Food Safety and Applied Nutrition /Food and Drug Administration /U.S. Department of Health and Human Services and Food Safety and Inspection Service/U.S. Department of Agriculture. (also available at https://www.fda.gov/Food/FoodScienceResearch/RiskSafetyAssessment/ucm183966.htm).
Chen, Y., Jackson, K.M., Chea, F.P. & Schaffner, D.W. 2001. Quantification and Variability Analysis of Bacterial Cross-Contamination Rates in Common Food Service Tasks. *Journal of Food Protection*, 64(1): 72–80. https://doi.org/10.4315/0362-028X-64.1.72
Chereau, F., Opatowski, L., Tourdjman, M. & Vong, S. 2017. Risk assessment for antibiotic resistance in South East Asia. *BMJ*: j3393. https://doi.org/10.1136/bmj.j3393
Christensen, B., Sommer, H., Rosenquist, H. & Nielsen, N. 2001. Risk Assessment on *Campylobacter jejuni* in Chicken Products. Danish Veterinary and Food Administration. (also available at http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.118.7218).
Cochran, W.G. 1950. Estimation of bacterial densities by means of the "most probable number". *Biometrics*, 6(2): 105–116.
Coleman, M. & Marks, H. 1998. Topics in Dose-Response Modeling. *Journal of Food Protection*, 61(11): 1550–1559. https://doi.org/10.4315/0362-028X-61.11.1550
Coleman, M.E., Marks, H.M., Golden, N.J. & Latimer, H.K. 2004. Discerning Strain Effects in Microbial Dose-Response Data. *Journal of Toxicology and Environmental Health, Part A*, 67(8–10): 667–685. https://doi.org/10.1080/15287390490428134
Coleman, M.E., Sandberg, S. & Anderson, S.A. 2003. Impact of Microbial Ecology of Meat and Poultry Products on Predictions from Exposure Assessment Scenarios for Refrigerated Storage. *Risk Analysis*, 23(1): 215–228. https://doi.org/10.1111/1539-6924.00301

Cox Jr., L.A. 2008. What's Wrong with Risk Matrices? *Risk Analysis*, 28(2): 497–512. https://doi.org/10.1111/j.1539-6924.2008.01030.x
Cox Jr., L.A., Babayev, D. & Huber, W. 2005. Some Limitations of Qualitative Risk Rating Systems. *Risk Analysis*, 25(3): 651–662. https://doi.org/10.1111/j.1539-6924.2005.00615.x
Cullen, A.C. 1994. Measures of Compounding Conservatism in Probabilistic Risk Assessment. *Risk Analysis*, 14(4): 389–393. https://doi.org/10.1111/j.1539-6924.1994.tb00257.x
Cullen, A.C. & Frey, H.C. 1999. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Springer US. (also available at http://www.springer.com/gp/book/9780306459566).
Dalgaard, P., Ross, T., Kamperman, L., Neumeyer, K. & McMeekin, T.A. 1994. Estimation of bacterial growth rates from turbidimetric and viable count data. *International Journal of Food Microbiology*, 23(3): 391–404. https://doi.org/10.1016/0168-1605(94)90165-1
De Knegt, L.V., Pires, S.M. & Hald, T. 2015. Attributing foodborne salmonellosis in humans to animal reservoirs in the European Union using a multi-country stochastic model. *Epidemiology & Infection*, 143(6): 1175–1186. https://doi.org/10.1017/S0950268814001903
DeDonder, S., Jacob, C.J., Surgeoner, B.V., Chapman, B., Phebus, R. & Powell, D.A. 2009. Self-reported and observed behavior of primary meal preparers and adolescents during preparation of frozen, uncooked, breaded chicken products. *British Food Journal*, 111(9): 915–929. https://doi.org/10.1108/00070700910992844
Dee, D.P. 1994. *Guidelines for documenting the validity of computational modelling software*. Delft, The Netherlands, IAHR/AIRH. 24 pp.
Dee, D.P. 1995. A Pragmatic Approach to Model Validation. *In* D.R. Lynch & A.M. Davies, eds. *Quantitative Skill Assessment for Coastal Ocean Models*, p. Coastal and Estuarine Studies. American Geophysical Union.
Dias, L.C., Morton, A. & Quigley, J., eds. 2018. *Elicitation*. International Series in Operations Research & Management Science. Springer International Publishing. (also available at http://link.springer.com/10.1007/978-3-319-65052-4).
Dogan, O.B., Clarke, J., Mattos, F. & Wang, B. 2019. A quantitative microbial risk assessment model of *Campylobacter* in broiler chickens: Evaluating processing interventions. *Food Control*, 100: 97–110. https://doi.org/10.1016/j.foodcont.2019.01.003
Dolan, K.D. & Mishra, D.K. 2013. Parameter Estimation in Food Science. *Annual Review of Food Science and Technology*, 4(1): 401–422. https://doi.org/10.1146/annurev-food-022811-101247
Dorny, P. & Praet, N. 2007. Taenia saginata in Europe. *Veterinary Parasitology*, 149(1–2): 22–24. https://doi.org/10.1016/j.vetpar.2007.07.004
Doucet, P. & Sloep, P.B. 1992. *Mathematical Modeling in the Life Sciences*. Ellis Horwood Series in Mathematics and Its Applications. New York, Ellis Horwood Ltd. 450 pp.
Duncan, G.E. 2014. Determining the health benefits of poultry industry compliance measures: the case of campylobacteriosis regulation in New Zealand. *The New Zealand Medical Journal*, 127(1391): 22–37.
Dziak, J.J., Coffman, D.L., Lanza, S.T., Li, R. & Jermiin, L.S. 2018. Sensitivity and Specificity of Information Criteria. *bioRxiv*: 449751. https://doi.org/10.1101/449751
Ebel, E.D. & Williams, M.S. 2019. Assessing the effectiveness of revised performance standards for *Salmonella* contamination of comminuted poultry. *Microbial Risk Analysis*: 100076. https://doi.org/10.1016/j.mran.2019.05.002
Ebel, E.D., Williams, M.S., Golden, N.J. & Marks, H.M. 2012. Simplified framework for predicting changes in public health from performance standards applied in slaughter establishments. *Food Control*, 28(2): 250–257. https://doi.org/10.1016/j.foodcont.2012.05.016
EcoSure. 2008. 2007 U.S. Cold Temperature Evaluation: Design and Summary Pages. EcoSure. (also available at http://foodrisk.org/exclusives/EcoSure/).

EFSA. 2004a. Statement on BSE/TSE and the health risks of the consumption of milk and milk derived products from goats by the Scientific Panel on biological hazards (BIOHAZ). *EFSA Journal*, 2(11). https://doi.org/10.2903/j.efsa.2004.136
EFSA. 2004b. EFSA publishes Geographical BSE-Risk (GBR) assessments for Australia, Canada, Mexico, Norway, South Africa, Sweden and the United States of America. In: *European Food Safety Authority* [online]. [Cited 12 October 2018]. https://www.efsa.europa.eu/en/press/news/040820
EFSA. 2012. Scientific Opinion on Risk Assessment Terminology. *EFSA Journal*, 10(5): 2664. https://doi.org/10.2903/j.efsa.2012.2664
EFSA. 2014a. Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment: Guidance on expert knowledge elicitation. *EFSA Journal*, 12(6): 3734. https://doi.org/10.2903/j.efsa.2014.3734
EFSA. 2014b. EFSA Panel on Biological Hazards (BIOHAZ) Scientific Opinion on the public health risks of table eggs due to deterioration and development of pathogens: Public health risks of table eggs due to deterioration and development of pathogens. *EFSA Journal*, 12(7): 3782. https://doi.org/10.2903/j.efsa.2014.3782
EFSA. 2018a. EFSA Scientific Committee: Guidance on Uncertainty Analysis in Scientific Assessments. *EFSA Journal*, 16(1): e05123. https://doi.org/10.2903/j.efsa.2018.5123
EFSA. 2018b. EFSA Scientific Committee: The principles and methods behind EFSA's Guidance on Uncertainty Analysis in Scientific Assessment. *EFSA Journal*, 16(1): e05122. https://doi.org/10.2903/j.efsa.2018.5122
Egorov, A.I., Griffin, S.M., Ward, H.D., Reilly, K., Fout, G.S. & Wade, T.J. 2018. Application of a salivary immunoassay in a prospective community study of waterborne infections. *Water Research*, 142: 289–300. https://doi.org/10.1016/j.watres.2018.05.030
Eschenbach, T.G. 1992. Spiderplots versus Tornado Diagrams for Sensitivity Analysis. *Interfaces*, 22(6): 40–46. https://doi.org/10.1287/inte.22.6.40
FAO. 2001. Joint FAO/WHO Expert Consultation on Risk Assessment of Microbiological Hazards in Foods: Risk characterization of *Salmonella* spp. in eggs and broiler chickens and *Listeria monocytogenes* in ready-to-eat foods. Food and Nutrition Paper No. 72. Rome, Italy, FAO [Food and Agricultural Organization of the United Nations]. (also available at http://www.fao.org/docrep/008/y1332e/y1332e00.htm).
FAO. 2017. Food Safety Risk Management: Evidence-informed policies and decisions, considering multiple factors. Food Safety and Quality No. 4. Rome, Italy, FAO [Food and Agricultural Organization of the United Nations]. (also available at http://www.fao.org/3/i8240en/I8240EN.pdf).
FAO/WHO. 2002a. Risk assessments of <i>Salmonella<i> in eggs and broiler chickens. Microbiological Risk Assessment Series No. 2. Rome, Italy, FAO/WHO. (also available at http://www.who.int/foodsafety/publications/salmonella/en/).
FAO/WHO. 2002b. Principles and guidelines for incorporating microbiological risk assessment in the development of food safety standards, guidelines and related texts. Report of a Joint FAO/WHO Consultation. Kiel, Germany, FAO/WHO. (also available at http://www.fao.org/3/a-y4302e.pdf).
FAO/WHO. 2003. Hazard Characterization for Pathogens in Food and Water. , p. 76. Microbiological Risk Assessment Series No. 3. Rome, Italy, FAO/WHO. (also available at http://www.who.int/foodsafety/publications/mra_3/en/).
FAO/WHO. 2004. Risk assessment of *Listeria monocytogenes* spp. in ready-to-eat foods. Microbiological Risk Assessment Series No. 5. Rome, Italy, FAO/WHO. (also available at http://www.who.int/foodsafety/publications/mra_listeria/en/).
FAO/WHO. 2005. Risk assessment of *Vibrio vulnificus* in raw oysters. Microbiological Risk Assessment Series No. 8. Rome, Italy, FAO/WHO. (also available at http://www.who.int/foodsafety/publications/mra8/en/).
FAO/WHO. 2006a. *Enterobacter sakazakii* and *Salmonella* in powdered infant formula. Microbiological Risk Assessment Series No. 10. Rome, Italy, FAO/WHO. (also available at https://www.who.int/foodsafety/publications/mra10/en/).

FAO/WHO. 2006b. The use of microbiological risk assessment outputs to develop practical risk management strategies: Metrics to improve food safety. Kiel, Germany, Food and Agriculture Organization of the United Nations/World Health Organization.
FAO/WHO. 2007. *Enterobacter sakazakii* and other microorganisms in powdered infant formula. Microbiological Risk Assessment Series No. 6. Rome, Italy, FAO/WHO. (also available at https://www.who.int/foodsafety/publications/mra6-enterobacter-sakazakii/en/).
FAO/WHO. 2008. Exposure assessment of microbiological hazards in food. Microbiological Risk Assessment Series No. 7. Rome, Italy, FAO/WHO. (also available at http://www.who.int/foodsafety/publications/mra7/en/).
FAO/WHO. 2009a. Risk characterization of microbiological hazards in food. Microbiological Risk Assessment Series No. 17. Rome, Italy, FAO/WHO. (also available at www.who.int/foodsafety/publications/risk-characterization/en/).
FAO/WHO. 2009b. *Salmonella* and *Campylobacter* in chicken meat: Meeting report. Microbiological Risk Assessment Series No. 19. Rome, Italy, FAO/WHO. (also available at http://www.who.int/foodsafety/publications/mra19/en/).
FAO/WHO. 2009c. Risk assessment of *Campylobacter* spp. in broiler chickens: Interpretative Summary. Microbiological Risk Assessment Series No. 11. Rome, Italy, FAO/WHO. (also available at http://www.fao.org/fileadmin/templates/agns/pdf/jemra/MRA_11.pdf).
FAO/WHO. 2009d. Risk assessment of *Campylobacter* spp. in broiler chickens. Microbiological Risk Assessment Series No. 12. Rome, Italy, FAO/WHO. (also available at http://www.fao.org/fileadmin/templates/agns/pdf/jemra/MRA_12.pdf).
FAO/WHO, ed. 2009e. *Principles and methods for the risk assessment of chemicals in food*. Environmental health criteria No. 240. Geneva, World Health Organization. 1 p. (also available at https://www.who.int/publications-detail/principles-and-methods-for-the-risk-assessment-of-chemicals-in-food).
FAO/WHO. 2011a. Risk assessment of *Vibrio parahaemolyticus* in raw and under-cooked seafood. Microbiological Risk Assessment Series No. 16. Rome, Italy, FAO/WHO. (also available at https://www.who.int/foodsafety/publications/mra_16/en/).
FAO/WHO. 2011b. Enterohaemorrhagic *Escherichia coli* in raw beef and beef products: approaches for the provision of scientific advice meeting report. , p. 126. Food and Agriculture Organization of the United Nations/World Health Organization.
FAO/WHO. 2014. Multicriteria-based ranking for risk management of food-borne parasites. Microbiological Risk Assessment Series No. 23. Rome, Italy, FAO/WHO. (also available at https://www.who.int/foodsafety/publications/mra_23/en/).
FAO/WHO. 2016. *Risk communication applied to food safety: Handbook*. Food safety and quality series No. 2. Rome, Italy, Food and Agricultural Organization of the United Nations, World Health Organization. (also available at http://www.who.int/foodsafety/Risk-Communication/en/).
FAO/WHO. 2018a. Shiga toxin-producing *Escherichia coli* (STEC) and food: attribution, characterization, and monitoring. Microbiological Risk Assessment Series No. 31. Rome, Italy, FAO/WHO. (also available at https://www.who.int/foodsafety/publications/mra_31/en/).
FAO/WHO. 2018b. *FAO/WHO Framework for the Provision of Scientific Advice on Food Safety and Nutrition*. Rome/Geneva, Food and Agricultural Organization of the United Nations, World Health Organization. (also available at http://www.fao.org/3/i7494en/I7494EN.pdf).
Fazil, A., Lowman, R., Stern, N. & Lammerding, A.M. 1999. Quantitative risk assessment model for *Campylobacter jejuni* in chicken. p. 65. Paper presented at 10th Int. Workshop on CHRO (Campylobacter, Helicobacter and Related Organisms), 1999, Baltimore, MD, USA.
Fazil, A.M. 2005. A primer on risk assessment modelling: focus on seafood products. Fisheries Technical Paper No. 462. Rome, Food and Agriculture Organization of the United Nations.
USFDA/USDA-FSIS. 2003. Quantitative Assessment of the relative risk to public health from foodborne *Listeria monocytogenes* among selected categories of ready-to-eat foods. Washington, DC, USA, USFDA/FSIS [U.S. Department of Health and Human Services - Food and Drug

Administration / U.S. Department of Agriculture - Food Safety and Inspection Service]. (also available at https://www.fda.gov/downloads/food/foodscienceresearch/ucm197330.pdf).
Frank, C., Milde-Busch, A. & Werber, D. 2014. Results of surveillance for infections with Shiga toxin-producing Escherichia coli (STEC) of serotype O104:H4 after the large outbreak in Germany, July to December 2011. *Euro Surveillance: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, 19(14).
Frey, H.C., Mokhtari, A. & Danish, T. 2003. Evaluation of Selected Sensitivity Analysis Methods Based Upon Applications to Two Food Safety Process Risk Models. Washington, DC, USA, USDA [Office of Risk Assessment and Cost-Benefit Analysis, U.S. Department of Agriculture]. (also available at https://www.ccee.ncsu.edu/wp-content/uploads/2015/08/risk-phase-2-final.pdf).
Frey, H.C., Mokhtari, A. & Zheng, J. 2004. Recommended Practice Regarding Selection, Application, and Interpretation of Sensitivity Analysis Methods Applied to Food Safety Process Risk Models. Washington, DC, USA, USDA [Office of Risk Assessment and Cost-Benefit Analysis, U.S. Department of Agriculture]. (also available at https://www.ccee.ncsu.edu/wp-content/uploads/2015/08/risk-phase-3-final.pdf).
Frey, H.C. & Patil, S.R. 2002. Identification and Review of Sensitivity Analysis Methods. *Risk Analysis*, 22(3): 553–578. https://doi.org/10.1111/0272-4332.00039
USDA-FSIS. 2001. Risk Assessment of the Public Health Impact of *Escherichia coli* O157:H7 in Ground Beef. USDA/FSIS [U.S. Department of Agriculture/Food Safety and Inspection Service]. (also available at https://www.fsis.usda.gov/wps/wcm/connect/1db13d79-1cd9-4e4d-b6ca-16ad89a085a1/00-023NReport.pdf?MOD=AJPERES).
USDA-FSIS. 2002. Comparative Risk Assessment for Intact (Non-Tenderized) and Non-Intact (Tenderized) Beef: Technical Report. Washington, DC, USA, USDA/FSIS [U.S. Department of Agriculture/Food Safety and Inspection Service]. (also available at https://www.fsis.usda.gov/shared/PDF/Beef_Risk_Assess_Report_Mar2002.pdf).
USDA-FSIS. 2005. Risk Assessments of *Salmonella* Enteritidis in Shell Eggs and *Salmonella* spp. in Egg Products. USDA/FSIS [U.S. Department of Agriculture/Food Safety and Inspection Service]. (also available at https://www.fsis.usda.gov/wps/wcm/connect/16abcb5f-21fc-4731-9346-fd8401582ba4/SE_Risk_Assess_Oct2005.pdf?MOD=AJPERES).
Funk, J.A., Davies, P.R. & Nichols, M.A. 2000. The effect of fecal sample weight on detection of *Salmonella* enterica in swine feces. *Journal of Veterinary Diagnostics*, 12: 412–418.
Gallagher, D., Pouillot, R., Hoelzer, K., Tang, J., Dennis, S.B. & Kause, J.R. 2016. *Listeria monocytogenes* in retail delicatessens: An interagency risk assessment—Risk mitigations. *Journal of Food Protection*, 79(7): 1076–1088. https://doi.org/10.4315/0362-028X.JFP-15-336
Gallagher, E., Ryan, J., Kelly, L., Leforban, Y. & Wooldridge, M. 2002. Estimating the risk of importation of foot-and-mouth disease into Europe | Veterinary Record. *Veterinary Record*, 150: 769–772. http://dx.doi.org/10.1136/vr.150.25.769
Gelman, A. & Carlin, J. 2014. Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, 9(6): 641–651. https://doi.org/10.1177/1745691614551642
Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. 2013. *Bayesian Data Analysis*. 3 edition edition. Boca Raton, Chapman and Hall/CRC. 675 pp.
van Gerwen, S.J.C., te Giffel, M.C., van't Riet, K., Beumer, R.R. & Zwietering, M.H. 2000. Stepwise quantitative risk assessment as a tool for characterization of microbiological food safety. *Journal of Applied Microbiology*, 88(6): 938–951. https://doi.org/10.1046/j.1365-2672.2000.01059.x
van Gerwen, S.J.C. & Gorris, L.G.M. 2004. Application of elements of microbiological risk assessment in the food industry via a tiered approach. *Journal of Food Protection*, 67(9): 2033–2040. https://doi.org/10.4315/0362-028X-67.9.2033
Giménez, B. & Dalgaard, P. 2004. Modelling and predicting the simultaneous growth of Listeria monocytogenes</i> and spoilage micro-organisms in cold-smoked salmon. *Journal of Applied Microbiology*, 96(1): 96–109. https://doi.org/10.1046/j.1365-2672.2003.02137.x

Gkogka, E., Reij, M.W., Gorris, L.G.M. & Zwietering, M.H. 2013a. The application of the Appropriate Level of Protection (ALOP) and Food Safety Objective (FSO) concepts in food safety management, using *Listeria monocytogenes* in deli meats as a case study. *Food Control*, 29(2): 382–393. https://doi.org/10.1016/j.foodcont.2012.04.020

Gkogka, E., Reij, M.W., Gorris, L.G.M. & Zwietering, M.H. 2013b. Risk assessment strategies as a tool in the application of the Appropriate Level of Protection (ALOP) and Food Safety Objective (FSO) by risk managers. *International Journal of Food Microbiology*, 167(1): 8–28. https://doi.org/10.1016/j.ijfoodmicro.2013.04.013

Gold, M.R., Stevenson, D. & Fryback, D.G. 2002. HALYs and QALYs and DALYs, Oh My: Similarities and Differences in Summary Measures of Population Health. *Annual Review of Public Health*, 23(1): 115–134. https://doi.org/10.1146/annurev.publhealth.23.100901.140513

Gonzales-Barron, U., Gonçalves-Tenório, A., Rodrigues, V. & Cadavez, V. 2017. Foodborne pathogens in raw milk and cheese of sheep and goat origin: a meta-analysis approach. *Current Opinion in Food Science*, 18: 7–13. https://doi.org/10.1016/j.cofs.2017.10.002

Gonzales-Barron, U., Kerr, M., Sheridan, J.J. & Butler, F. 2010. Count data distributions and their zero-modified equivalents as a framework for modelling microbial data with a relatively high occurrence of zero counts. *International Journal of Food Microbiology*, 136(3): 268–277.

Gonzales-Barron, U., Piza, L., Xavier, C., Costa, E. & Cadavez, V. 2016. An exposure assessment model of the prevalence of *Salmonella* spp. along the processing stages of Brazilian beef. *Food Science and Technology International*, 22(1): 10–20. https://doi.org/10.1177/1082013214560446

Greig, J.D. & Ravel, A. 2009. Analysis of foodborne outbreak data reported internationally for source attribution. *International Journal of Food Microbiology*, 130(2): 77–87. https://doi.org/10.1016/j.ijfoodmicro.2008.12.031

Greppi, A. & Rantsiou, K. 2016. Methodological advancements in foodborne pathogen determination: from presence to behavior. *Current Opinion in Food Science*, 8: 80–88. https://doi.org/10.1016/j.cofs.2016.04.005

Guo, C., Hoekstra, R.M., Schroeder, C.M., Pires, S.M., Ong, K.L., Hartnett, E., Naugle, A., Harman, J., Bennett, P., Cieslak, P., Scallan, E., Rose, B., Holt, K.G., Kissler, B., Mbandi, E., Roodsari, R., Angulo, F.J. & Cole, D. 2011. Application of Bayesian Techniques to Model the Burden of Human Salmonellosis Attributable to U.S. Food Commodities at the Point of Processing: Adaptation of a Danish Model. *Foodborne Pathogens and Disease*, 8(4): 509–516. https://doi.org/10.1089/fpd.2010.0714

Guo, M., Buchanan, R.L., Dubey, J.P., Hill, D.E., Lambertini, E., Ying, Y., Gamble, H.R., Jones, J.L. & Pradhan, A.K. 2015. Qualitative Assessment for *Toxoplasma gondii* Exposure Risk Associated with Meat Products in the United States. *Journal of Food Protection*, 78(12): 2207–2219. https://doi.org/10.4315/0362-028X.JFP-15-270

Gustavsson, J., Cederberg, C. & Sonesson, U. 2011. Global Food Losses and Food Waste. Paper presented at Save Food Congress, 16 May 2011, Duesseldorf, Germany. [Cited 31 March 2020]. https://www.researchgate.net/publication/267919405_Global_Food_Losses_and_Food_Waste

Haas, C.N. 1983. Estimation of risk due to low doses of microorganisms: A comparison of alternative methodolodies. *American Journal of Epidemiology*, 118(4): 573–582.

Haas, C.N. 2002. Conditional Dose-Response Relationships for Microorganisms: Development and Application. *Risk Analysis*, 22(3): 455–463. https://doi.org/10.1111/0272-4332.00035

Haas, C.N. 2015. Microbial Dose Response Modeling: Past, Present, and Future. *Environmental Science & Technology*, 49(3): 1245–1259. https://doi.org/10.1021/es504422q

Haas, C.N., Rose, J.B. & Gerba, C.P. 2014. *Quantitative Microbial Risk Assessment*. 2nd edition. Wiley. 440 pp. (also available at https://www.wiley.com/en-au/Quantitative+Microbial+Risk+Assessment%2C+2nd+Edition-p-9781118910023).

Haddad, N., Johnson, N., Kathariou, S., Métris, A., Phister, T., Pielaat, A., Tassou, C., Wells-Bennik, M.H.J. & Zwietering, M.H. 2018. Next generation microbiological risk assessment—Potential of

omics data for hazard characterisation. *International Journal of Food Microbiology*, 287: 28–39. https://doi.org/10.1016/j.ijfoodmicro.2018.04.015
Hald, T., Vose, D., Wegener, H.C. & Koupeev, T. 2004. A Bayesian Approach to Quantify the Contribution of Animal-Food Sources to Human Salmonellosis. *Risk Analysis*, 24(1): 255–269. https://doi.org/10.1111/j.0272-4332.2004.00427.x
Hald, T., Wingstrand, A., Brøndsted, T. & Wong, D.M.A.L.F. 2006. Human Health Impact of *Salmonella* Contamination in Imported Soybean Products: A Semiquantitative Risk Assessment. *Foodborne Pathogens and Disease*, 3(4): 422–431. https://doi.org/10.1089/fpd.2006.3.422
Hall, G., Keflemariam, Y., Raupach, J., Becker, N. & Kirk, M. 2008. Estimating Community Incidence of *Salmonella*, *Campylobacter*, and Shiga Toxin–producing *Escherichia coli* Infections, Australia. *Emerging Infectious Diseases*, 14(10): 1601–1609. https://doi.org/10.3201/eid1410.071042
Hansen, L.P. 1982. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50(4): 1029–1054. https://doi.org/10.2307/1912775
Hartnett, E. 2002. Campylobacter *in broilers: a quantitative risk assessment approach*. Glasgow, United Kingdom, University of Strathclyde. (PhD Thesis)
Hartnett, E., Kelly, L., Newell, D., Wooldridge, M. & Gettinby, G. 2001. A quantitative risk assessment for the occurrence of *Campylobacter* in chickens at the point of slaughter. *Epidemiology & Infection*, 127(2): 195–206. https://doi.org/10.1017/S0950268801005866
Havelaar, A.H., de Wit, M.A.S., van Koningsveld, R. & van Kempen, E. 2000. Health burden in the Netherlands due to infection with thermophilic *Campylobacter* spp. *Epidemiology & Infection*, 125(3): 505–522.
Hayama, Y., Yamamoto, T., Kasuga, F. & Tsutsui, T. 2011. Simulation Model for *Campylobacter* Cross-Contamination During Poultry Processing at Slaughterhouses. *Zoonoses and Public Health*, 58(6): 399–406. https://doi.org/10.1111/j.1863-2378.2010.01385.x
Hemming, V., Burgman, M.A., Hanea, A.M., McBride, M.F. & Wintle, B.C. 2018. A practical guide to structured expert elicitation using the IDEA protocol. *Methods in Ecology and Evolution*, 9(1): 169–180. https://doi.org/10.1111/2041-210X.12857
Holcomb, D.L., Smith, M.A., Ware, G.O., Hung, Y.-C., Brackett, R.E. & Doyle, M.P. 1999. Comparison of six dose-response models for use with food-borne pathogens. *Risk Analysis*, 19(6): 1091–1100.
Hsi, D.J., Ebel, E.D., Williams, M.S., Golden, N.J. & Schlosser, W.D. 2015. Comparing foodborne illness risks among meat commodities in the United States. *Food Control*, 54: 353–359. https://doi.org/10.1016/j.foodcont.2015.02.018
Huang, L. 2017. IPMP Global Fit – A one-step direct data analysis tool for predictive microbiology. *International Journal of Food Microbiology*, 262: 38–48. https://doi.org/10.1016/j.ijfoodmicro.2017.09.010
ICMSF. 1996. *Microorganisms in Foods 5: Characteristics of Microbial Pathogens*. Kluwer Academic / Plenum Publishers.
ILSI. 2000. *Revised framework for microbial risk assessment.* Washington, DC, International Life Science Institute Press. (also available at https://archive.epa.gov/raf/web/pdf/ilsi-ra-framework.pdf).
IPCC. 2001. Good practice guidance and uncertainty management in national greenhouse gas inventories. Intergovernmental Panel on Climate Change. (also available at https://www.ipcc-nggip.iges.or.jp/public/gp/english/gpgaum_en.html).
Jameson, J.E. 1962. A discussion of the dynamics of *Salmonella* enrichment. *Epidemiology & Infection*, 60(2): 193–207. https://doi.org/10.1017/S0022172400039462
Jaykus, L.A. 1996. The Application of Quantitative Risk Assessment to Microbial Food Safety Risks. *Critical Reviews in Microbiology*, 22(4): 279–293. https://doi.org/10.3109/10408419609105483
Jenkinson, D. 2005. The elicitation of probabilities: A review of the statistical literature. University of Sheffield.

Jensen, D.A., Friedrich, L.M., Harris, L.J., Danyluk, M.D. & Schaffner, D.W. 2015. Cross contamination of *Escherichia coli* O157:H7 between lettuce and wash water during home-scale washing. *Food Microbiology*, 46: 428–433. https://doi.org/10.1016/j.fm.2014.08.025
Jewell, K. 2012. Comparison of 1-step and 2-step methods of fitting microbiological models. *International Journal of Food Microbiology*, 160(2): 145–161.
Johnson, N.L., A. W. Kemp & Kotz, S. 2008. *Continuous Univariate Distributions*. 3rd Edition edition. Wiley.
Jongenburger, I., Bassett, J., Jackson, T., Zwietering, M.H. & Jewell, K. 2012. Impact of microbial distributions on food safety I. Factors influencing microbial distributions and modelling aspects. *Food Control*, 26(2): 601–609.
Jongenburger, I., Reij, M.W., Boer, E.P.J., Gorris, L.G.M. & Zwietering, M.H. 2011. Actual distribution of *Cronobacter*spp. in industrial batches of powdered infant formula and consequences for performance of sampling strategies. *International Journal of Food Microbiology*, 151(1): 62–69.
Kaplan, S. & Garrick, B.J. 1981. On The Quantitative Definition of Risk. *Risk Analysis*, 1(1): 11–27. https://doi.org/10.1111/j.1539-6924.1981.tb01350.x
Kass, R.E. & Raftery, A.E. 1995. Bayes Factors. *Journal of the American Statistical Association*, 90(430): 773–795. https://doi.org/10.1080/01621459.1995.10476572
Kasuga, F., Hirota, M., Wada, M., Yunokawa, T., Toyofuku, H., Shibatsuji, M., Michino, H., Kuwasaki, T., Yamamoto, S. & Kumagai, S. 2004. Archiving of Food Samples from Restaurants and Caterers—Quantitative Profiling of Outbreaks of Foodborne Salmonellosis in Japan. *Journal of Food Protection*, 67(9): 2024–2032. https://doi.org/10.4315/0362-028X-67.9.2024
Kiermeier, A., Jenson, I. & Sumner, J. 2014. Risk assessment of *Escherichia coli* O157 illness from consumption of hamburgers in the USA made from Australian manufacturing beef. *Risk Analysis*.
Kiermeier, A., Jenson, I. & Sumner, J. 2015. Risk Assessment of *Escherichia coli* O157 Illness from Consumption of Hamburgers in the United States Made from Australian Manufacturing Beef: *E. coli* O157 Illness from Consumption of Hamburgers. *Risk Analysis*, 35(1): 77–89. https://doi.org/10.1111/risa.12248
Kiermeier, A., Mellor, G., Barlow, R. & Jenson, I. 2011. Assumptions of Acceptance Sampling and the Implications for Lot Contamination: \emphE. coli O157 in Lots of Australian Manufacturing Beef. *Journal of Food Protection*, 74(4): 539–544.
Kim, S.A., Yun, S.J., Lee, S.H., Hwang, I.G. & Rhee, M.S. 2013. Temperature increase of foods in car trunk and the potential hazard for microbial growth. *Food Control*, 29: 66–70.
King, N., Lake, R. & Cressey, P. 2011. Risk profile: *Salmonella* (non-typhoid) in poultry (whole and pieces). No. FW11044. ESR [Environmental Sciences and Research Limited]. (also available at https://www.mpi.govt.nz/food-safety/food-safety-and-suitability-research/food-risk-assessment/food-risk-profiles/).
Knuth, D.E. 1992. *Literate programming*. CSLI lecture notes No. no. 27. Stanford, Calif., Center for the Study of Language and Information. 368 pp.
Kosa, K.M., Cates, S.C., Bradley, S., Chambers, E. & Godwin, S. 2015. Consumer-Reported Handling of Raw Poultry Products at Home: Results from a National Survey. *Journal of Food Protection*, 78(1): 180–186. https://doi.org/10.4315/0362-028X.JFP-14-231
Koutsoumanis, K.P., Lianou, A. & Gougouli, M. 2016. Latest developments in foodborne pathogens modeling. *Current Opinion in Food Science*, 8: 89–98. https://doi.org/10.1016/j.cofs.2016.04.006
Kruschke, J. 2014. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. 2 edition edition. Boston, Academic Press. 776 pp.
Lake, R. & Cressey, P. 2013. Risk profile: *Campylobacter jejuni/coli* in poultry (whole and pieces). No. FW11065. ESR [Environmental Sciences and Research Limited]. (also available at https://www.mpi.govt.nz/food-safety/food-safety-and-suitability-research/food-risk-assessment/food-risk-profiles/).
Lake, R., Gilbert, S., Wong, T.L. & Cressey, P. 2009. Risk profile: *Mycobacterium bovis* (non-typhoid) in milk. No. FW08031. ESR [Environmental Sciences and Research Limited]. (also available at

https://www.mpi.govt.nz/food-safety/food-safety-and-suitability-research/food-risk-assessment/food-risk-profiles/).
Lake, R., Hudson, A., Cressey, P. & Gilbert, S. 2005. Risk profile: *Listeria Monocytogenes* in ready-to-eat salads. No. FW0446. ESR [Environmental Sciences and Research Limited]. (also available at https://www.foodsafety.govt.nz/elibrary/industry/Risk_Profile_Listeria_Monocytogenes_Ready-Science_Research.pdf).
Lake, R.J., Cressey, P.J., Campbell, D.M. & Oakley, E. 2010. Risk Ranking for Foodborne Microbial Hazards in New Zealand: Burden of Disease Estimates. *Risk Analysis*, 30(5): 743–752. https://doi.org/10.1111/j.1539-6924.2009.01269.x
Law, A.M. 2014. *Simulation Modeling and Analysis*. Fifth edition. Dubuque, McGraw-Hill Education. 800 pp.
Lawson, L.G., Jensen, J.D., Christiansen, P. & Lund, M. 2009. Cost-effectiveness of *Salmonella* reduction in Danish abattoirs. *International Journal of Food Microbiology*, 134(1–2): 126–132. https://doi.org/10.1016/j.ijfoodmicro.2009.03.024
Laxminarayan, R. & Macauley, M.K., eds. 2012. *The Value of Information*. Dordrecht, Springer Netherlands. (also available at http://link.springer.com/10.1007/978-94-007-4839-2).
Le Marc, Y., Valík, L. & Medveďová, A. 2009. Modelling the effect of the starter culture on the growth of Staphylococcus aureus in milk. *International Journal of Food Microbiology*, 129(3): 306–311. https://doi.org/10.1016/j.ijfoodmicro.2008.12.015
Lefkowitz, A., Fout, G.S., Losonsky, G., Wasserman, S.S., Israel, E. & Morris Jr., J.G. 1992. A Serosurvey of Pathogens Associated with Shellfish: Prevalence of Antibodies to *Vibrio* Species and Norwalk Virus in the Chesapeake Bay Region. *American Journal of Epidemiology*, 135(4): 369–380. https://doi.org/10.1093/oxfordjournals.aje.a116298
Lever, J., Krzywinski, M. & Altman, N. 2016. Model selection and overfitting. *Nature Methods*, 13(9): 703–704. https://doi.org/10.1038/nmeth.3968
Levine, E.S. 2012. Improving risk matrices: the advantages of logarithmically scaled axes. *Journal of Risk Research*, 15(2): 209–222. https://doi.org/10.1080/13669877.2011.634514
Linstone, H.A. & Turoff, M., eds. 2002. *The Delphi Method: Techniques and Application*. Second edition. (also available at https://web.njit.edu/~turoff/pubs/delphibook/index.html).
Lorimer, M.F. & Kiermeier, A. 2007. Analysing microbiological data: Tobit or not Tobit? *International Journal of Food Microbiology*, 116(3): 313–318.
Mangen, M.-J.J., Batz, M.B., Käsbohrer, A., Hald, T., Morris, J.G., Taylor, M. & Havelaar, A.H. 2010. Integrated Approaches for the Public Health Prioritization of Foodborne and Zoonotic Pathogens. *Risk Analysis*, 30(5): 782–797. https://doi.org/10.1111/j.1539-6924.2009.01291.x
Mangen, M.-J.J., Bouwknegt, M., Friesema, I.H.M., Haagsma, J.A., Kortbeek, L.M., Tariq, L., Wilson, M., van Pelt, W. & Havelaar, A.H. 2015. Cost-of-illness and disease burden of food-related pathogens in the Netherlands, 2011. *International Journal of Food Microbiology*, 196: 84–93. https://doi.org/10.1016/j.ijfoodmicro.2014.11.022
Mariner, J.C. & Paskin, R. 2000. Manual on participatory epidemiology: methods for the collection of action-oriented epidemiological intelligence. *Manual on participatory epidemiology: methods for the collection of action-oriented epidemiological intelligence.* (also available at https://www.cabdirect.org/cabdirect/abstract/20026790917).
Marks, H.M., Coleman, M.E., Lin, C.-T.J. & Roberts, T. 1998. Topics in Microbial Risk Assessment: Dynamic Flow Tree Process. *Risk Analysis*, 18(3): 309–328. https://doi.org/10.1111/j.1539-6924.1998.tb01298.x
McCullagh, P. & Nelder, J.A. 1989. *Generalized Linear Models*. Chapman and Hall.
McKellar, R.J. & Lu, X., eds. 2003. *Modeling Microbial Responses in Food*. First edition. Contemporary Food Science. CRC Press. (also available at https://www.crcpress.com/Modeling-Microbial-Responses-in-Food/McKellar-Lu/p/book/9780849312373).

McLinden, T., Sargeant, J.M., Thomas, M.K., Papadopoulos, A. & Fazil, A. 2014. Component costs of foodborne illness: a scoping review. *BMC Public Health*, 14(1): 509. https://doi.org/10.1186/1471-2458-14-509

McMeekin, T., Bowman, J., McQuestin, O., Mellefont, L., Ross, T. & Tamplin, M. 2008. The future of predictive microbiology: Strategic research, innovative applications and great expectations. *International Journal of Food Microbiology*, 128(1): 2–9. https://doi.org/10.1016/j.ijfoodmicro.2008.06.026

McMeekin, T.A., Olley, J., Ross, T. & Ratkowsky, D.A. 1993. *Predictive microbiology : theory and application*. Innovation in Microbiology No. 5. Taunton : Research studies press. (also available at http://lib.ugent.be/catalog/rug01:001474154).

McQuestin, O.J., Shadbolt, C.T. & Ross, T. 2009. Quantification of the Relative Effects of Temperature, pH, and Water Activity on Inactivation of *Escherichia coli* in Fermented Meat by Meta-Analysis. *Applied and Environmental Microbiology*, 75(22): 6963–6972. https://doi.org/10.1128/AEM.00291-09

Mellefont, L.A., McMeekin, T.A. & Ross, T. 2008. Effect of relative inoculum concentration on *Listeria monocytogenes* growth in co-culture. *International Journal of Food Microbiology*, 121(2): 157–168. https://doi.org/10.1016/j.ijfoodmicro.2007.10.010

Membré, J.-M. & Boué, G. 2018. Quantitative microbiological risk assessment in food industry: Theory and practical application. *Food Research International*, 106: 1132–1139. https://doi.org/10.1016/j.foodres.2017.11.025

Merrell, D.S., Butler, S.M., Qadri, F., Dolganov, N.A., Alam, A., Cohen, M.B., Calderwood, S.B., Schoolnik, G.K. & Camilli, A. 2002. Host-induced epidemic spread of the cholera bacterium. *Nature*, 417(6889): 642–645. https://doi.org/10.1038/nature00778

Messner, M.J. & Berger, P. 2016. Cryptosporidium Infection Risk: Results of New Dose-Response Modeling. *Risk Analysis*, 36(10): 1969–1982. https://doi.org/10.1111/risa.12541

Mokhtari, A., Frey, H.C. & Jaykus, L.-A. 2006. Application of classification and regression trees for sensitivity analysis of the *Escherichia coli* O157:H7 food safety process risk model. *Journal of Food Protection*, 69(3): 609–618.

Monge, S., Pijnacker, R., van Pelt, W., Franz, E., Kortbeek, L.M. & Mangen, M.-J.J. 2019. Accounting for long-term manifestations of *Cryptosporidium* spp infection in burden of disease and cost-of-illness estimations, the Netherlands (2013–2017). *PLOS ONE*, 14(3): e0213752. https://doi.org/10.1371/journal.pone.0213752

Moore, C.M., Sheldon, B.W. & Jaykus, L.-A. 2003. Transfer of *Salmonella* and *Campylobacter* from Stainless Steel to Romaine Lettuce. *Journal of Food Protection*, 66(12): 2231–2236. https://doi.org/10.4315/0362-028X-66.12.2231

Morgan, M.G., Henrion, M. & Small, M. 1992. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press. 354 pp. (also available at https://www.cambridge.org/gb/academic/subjects/psychology/cognition/uncertainty-guide-dealing-uncertainty-quantitative-risk-and-policy-analysis?format=HB&isbn=9780521365420).

Mullner, P., Spencer, S.E.F., Wilson, D.J., Jones, G., Noble, A.D., Midwinter, A.C., Collins-Emerson, J.M., Carter, P., Hathaway, S. & French, N.P. 2009. Assigning the source of human campylobacteriosis in New Zealand: A comparative genetic and epidemiological approach. *Infection, Genetics and Evolution*, 9(6): 1311–1319. https://doi.org/10.1016/j.meegid.2009.09.003

Murray, R., Glass-Kaastra, S., Gardhouse, C., Marshall, B., Ciampa, N., Franklin, K., Hurst, M., Thomas, M.K. & Nesbitt, A. 2017. Canadian Consumer Food Safety Practices and Knowledge: Foodbook Study. *Journal of Food Protection*, 80(10): 1711–1718. https://doi.org/10.4315/0362-028X.JFP-17-108

Mylius, S.D., Nauta, M.J. & Havelaar, A.H. 2007. Cross-Contamination During Food Preparation: A Mechanistic Model Applied to Chicken-Borne *Campylobacter*. *Risk Analysis*, 27(4): 803–813. https://doi.org/10.1111/j.1539-6924.2006.00872.x

NACMCF. 2010. Parameters for Determining Inoculated Pack/Challenge Study Protocols. *Journal of Food Protection*, 73(1): 140–202. https://doi.org/10.4315/0362-028X-73.1.140
NAS. 2002. Escherichia coli *O157:H7 in Ground Beef: Review of a Draft Risk Assessment*. Washington, DC, USA, National Academies Press. (also available at https://www.nap.edu/catalog/10528/escherichia-coli-o157h7-in-ground-beef-review-of-a-draft).
Nauta, M.J. 2000. Separation of uncertainty and variability in quantitative microbial risk assessment models. *International Journal of Food Microbiology*, 57(1): 9–18. https://doi.org/10.1016/S0168-1605(00)00225-7
Nauta, M.J. 2001. A modular process risk model structure for quantiative microbiolgical risk assessment and its application in an exposure assessment of *Bacillus cereus* in a REPFED. , p. 100. RIVM [Rijksinstituut voor Volksgezondheid en Milieu]. (also available at www.rivm.nl/bibliotheek/rapporten/149106007.html).
Nauta, M.J. 2008. The Modular Process Risk Model (MPRM): a Structured Approach to Food Chain Exposure Assessment (Ch 4). *In* D.W. Schaffner, ed. *Microbial Risk Analysis of Foods*, pp. 99–136. (also available at http://www.asmscience.org/content/book/10.1128/9781555815752.ch04).
Nauta, M.J., Evers, E.G., Takumi, K. & Havelaar, A.H. 2001. Risk assessment of Shiga-toxin producing *Escherichia coli* O157 in steak tartare in The Netherlands. , p. 169. Bilthoven, The Netherlands, RIVM [Rijksinstituut voor Volksgezondheid en Milieu]. (also available at https://www.rivm.nl/bibliotheek/rapporten/257851003.html).
Nauta, M.J., Van der Fels-Klerx, I. & Havelaar, A. 2005. A Poultry-Processing Model for Quantitative Microbiological Risk Assessment. *Risk Analysis*, 25(1): 85–98. https://doi.org/10.1111/j.0272-4332.2005.00569.x
Nauta, M.J., van der Wal, F.J., Putirulan, F.F., Post, J., van de Kassteele, J. & Bolder, N.M. 2009. Evaluation of the “testing and scheduling” strategy for control of \emphCampylobacter in broiler meat in The Netherlands. *International Journal of Food Microbiology*, 134: 216–222.
NHMRC. 2018. *Australian Drinking Water Guidelines 6 (2011; Update 2018)*. 3.5 edition. National Health and Medical Research Council. (also available at https://nhmrc.gov.au/about-us/publications/australian-drinking-water-guidelines#block-views-block-file-attachments-content-block-1).
Notermans, S., Dufrenne, J., Teunis, P. & Chackraborty, T. 1998. Studies on the Risk Assessment of *Listeria monocytogenes*. *Journal of Food Protection*, 61(2): 244–248. https://doi.org/10.4315/0362-028X-61.2.244
NRC. 1989. *Biologic Markers in Reproductive Toxicology*. Washington, D.C., National Academies Press. (also available at https://doi.org/10.17226/774).
OIE. 2018. *Terrestial Animal Health Code*. 27th edition. World Organization for Animal Health. (also available at http://www.oie.int/standard-setting/terrestrial-code/access-online/).
Omurtag, I., Paulsen, P., Hilbert, F. & Smulders, F.J.M. 2013. The risk of transfer of foodborne bacterial hazards in Turkey through the consumption of meat; risk ranking of muscle foods with the potential to transfer *Campylobacter* spp. *Food Security*, 5(1): 117–127. https://doi.org/10.1007/s12571-012-0230-z
Painter, J.A., Hoekstra, R.M., Ayers, T., Tauxe, R.V., Braden, C.R., Angulo, F.J. & Griffin, P.M. 2013. Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998–2008. *Emerging Infectious Diseases*, 19(3): 407–415.
Pang, H., Lambertini, E., Buchanan, R.L., Schaffner, D.W. & Pradhan, A.K. 2017. Quantitative Microbial Risk Assessment for Escherichia coli O157:H7 in Fresh-Cut Lettuce. *Journal of Food Protection*, 80(2): 302–311. https://doi.org/10.4315/0362-028X.JFP-16-246
Pérez-Rodríguez, F., Campos, D., Ryser, E.T., Buchholz, A.L., Posada-Izquierdo, G.D., Marks, B.P., Zurera, G. & Todd, E. 2011. A mathematical risk model for *Escherichia coli* O157:H7 cross-contamination of lettuce during processing. *Food Microbiology*, 28(4): 694–701.
Pérez-Rodríguez, F. & Valero, A. 2013. *Predictive microbiology in foods*. SpringerBriefs in food, health, and nutrition No. 5. New York, Springer. 128 pp.

Pérez-Rodríguez, F., Valero, A., Carrasco, E., García, R.M. & Zurera, G. 2008. Understanding and modelling bacterial transfer to foods: a review. *Trends in Food Science & Technology*, 19(3): 131–144. https://doi.org/10.1016/j.tifs.2007.08.003

Petitti, D.B. 2000. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine*. 2 edition edition. New York, Oxford University Press. 306 pp.

Pichler, J., Ziegler, J., Aldrian, U. & Allerberger, F. 2014. Evaluating levels of knowledge on food safety among food handlers from restaurants and various catering businesses in Vienna, Austria 2011/2012. *Food Control*, 35(1): 33–40. https://doi.org/10.1016/j.foodcont.2013.06.034

Pires, S.M. 2013. Assessing the applicability of currently available methods for attributing foodborne disease to sources, including food and food commodities. *Foodborne Pathogens and Disease*, 10(3): 206–13.

Pires, S.M., Evers, E.G., van Pelt, W., Ayers, T., Scallan, E., Angulo, F.J., Havelaar, A. & Hald, T. 2009. Attributing the Human Disease Burden of Foodborne Infections to Specific Sources. *Foodborne Pathogens and Disease*, 6(4): 417–424. https://doi.org/10.1089/fpd.2008.0208

Possas, A., Carrasco, E., García-Gimeno, R.M. & Valero, A. 2017. Models of microbial cross-contamination dynamics. *Current Opinion in Food Science*, 14: 43–49. https://doi.org/10.1016/j.cofs.2017.01.006

Pouillot, R., Chen, Y. & Hoelzer, K. 2015. Modeling number of bacteria per food unit in comparison to bacterial concentration in quantitative risk assessment: Impact on risk estimates. *Food Microbiology*, 45: 245–253. https://doi.org/10.1016/j.fm.2014.05.008

Pouillot, R. & Delignette-Muller, M.L. 2010. Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages. *International Journal of Food Microbiology*, 142(3): 330–340.

Pouillot, R., Gallagher, D., Tang, J., Hoelzer, K., Kause, J. & Dennis, S.B. 2015a. Listeria monocytogenes in Retail Delicatessens: An Interagency Risk Assessment—Model and Baseline Results. *Journal of Food Protection*, 78(1): 134–145. https://doi.org/10.4315/0362-028X.JFP-14-235

Pouillot, R., Garin, B., Ravaonindrina, N., Diop, K., Ratsitorahina, M., Ramanantsoa, D. & Rocourt, J. 2012. A Risk Assessment of Campylobacteriosis and Salmonellosis Linked to Chicken Meals Prepared in Households in Dakar, Senegal. *Risk Analysis*, 32(10): 1798–1819.

Pouillot, R., Hoelzer, K., Chen, Y. & Dennis, S.B. 2015b. *Listeria monocytogenes* Dose Response Revisited—Incorporating Adjustments for Variability in Strain Virulence and Host Susceptibility. *Risk Analysis*, 35(1): 90–108. https://doi.org/10.1111/risa.12235

Pouillot, R., Klontz, K.C., Chen, Y., Burall, L.S., Macarisin, D., Doyle, M., Bally, K.M., Strain, E., Datta, A.R., Hammack, T.S. & Van Doren, J.M. 2016. Infectious Dose of *Listeria monocytogenes* in Outbreak Linked to Ice Cream, United States, 2015. *Emerging Infectious Diseases*, 22(12): 2113–2119. https://doi.org/10.3201/eid2212.160165

Pouillot, R., Miconnet, N., Afchain, A.-L., Delignette-Muller, M.L., Beaufort, A., Rosso, L., J.-B, D. & Cornu, M. 2007. Quantitative Risk Assessment of \emphListeria monocytogenes in French Cold-Smoked Salmon: I. Quantitative Exposure Assessment. *Risk Analysis*, 27(3): 683–700. https://doi.org/10.1111/j.1539-6924.2007.00921.x

Powell, M., Ebel, E. & Schlosser, W. 2001. Considering uncertainty in comparing the burden of illness due to foodborne microbial pathogens. *International Journal of Food Microbiology*, 69(3): 209–215. https://doi.org/10.1016/S0168-1605(01)00495-0

Powell, M., Schlosser, W. & Ebel, E. 2004. Considering the complexity of microbial community dynamics in food safety risk assessment. *International Journal of Food Microbiology*, 90(2): 171–179. https://doi.org/10.1016/S0168-1605(03)00106-5

Prado-Silva, L., Cadavez, V., Gonzales-Barron, U., Rezende, A.C.B. & Sant'Ana, A.S. 2015. Meta-analysis of the Effects of Sanitizing Treatments on *Salmonella*, *Escherichia coli* O157:H7, and *Listeria monocytogenes* Inactivation in Fresh Produce. *Applied and Environmental Microbiology*, 81(23): 8008–8021. https://doi.org/10.1128/AEM.02216-15

Pressman, R.S. 2005. *Software Engineering: A Practitioner's Approach*. Palgrave Macmillan. 922 pp.

Pujol, L., Albert, I., Johnson, N.B. & Membré, J.-M. 2013. Potential application of quantitative microbiological risk assessment techniques to an aseptic-UHT process in the food industry. *International Journal of Food Microbiology*, 162(3): 283–296. https://doi.org/10.1016/j.ijfoodmicro.2013.01.021

Rantsiou, K., Kathariou, S., Winkler, A., Skandamis, P., Saint-Cyr, M.J., Rouzeau-Szynalski, K. & Amézquita, A. 2018. Next generation microbiological risk assessment: opportunities of whole genome sequencing (WGS) for foodborne pathogen surveillance, source tracking and risk assessment. *International Journal of Food Microbiology*, 287: 3–9. https://doi.org/10.1016/j.ijfoodmicro.2017.11.007

Ratkowsky, D.A. 1992. Predicting response times in predictive food microbiology. No. Occasional Paper No 1992/1. Hobart, Australia, Department of Primary Industry, Fisheries and Energy (Tasmania).

Ratkowsky, D.A., Ross, T., Macario, N., Dommett, T.W. & Kamperman, L. 1996. Choosing probability distributions for modelling generation time variability. *Journal of Applied Bacteriology*, 80(2): 131–137. https://doi.org/10.1111/j.1365-2672.1996.tb03200.x

Ratkowsky, D.A., Ross, T., McMeekin, T.A. & Olley, J. 1991. Comparison of Arrhenius-type and Bêlehrádek-type models for prediction of bacterial growth in foods. *Journal of Applied Bacteriology*, 71(5): 452–459. https://doi.org/10.1111/j.1365-2672.1991.tb03816.x

Roberts, T., Ahl, A. & McDowell, R. 1995. Risk assessment for foodborne microbial hazards. *Tracking foodborne pathogens from farm to table*, p. Miscellaneous Publications No. 1532. Washington, DC, USA, U.S. Department of Agriculture.

Ross, T. 2008. Microbial Ecology in Food Safety Risk Assessment (Ch 3). *In* D.W. Schaffner, ed. *Microbial Risk Analysis of Foods*, pp. 51–97. (also available at http://www.asmscience.org/content/book/10.1128/9781555815752.ch03).

Ross, T. & Dalgaard, P. 2004. Secondary models (Ch 4). *In* R.C. McKellar & X. Lu, eds. *Modeling microbial responses in food*, p. 360. Contemporary Food Science

Ross, T., Dalgaard, P. & Tienungoon, S. 2000. Predictive modelling of the growth and survival of Listeria in fishery products. *International Journal of Food Microbiology*, 62(3): 231–245. https://doi.org/10.1016/S0168-1605(00)00340-8

Ross, T. & McMeekin, T.A. 2003. Modeling Microbial Growth Within Food Safety Risk Assessments. *Risk Analysis*, 23(1): 179–197. https://doi.org/10.1111/1539-6924.00299

Ross, T., McMeekin, T.A. & Baranyi, J. 2014. Predictive Microbiology and Food Safety. *Encyclopedia of Food Microbiology*. 2nd edition, pp. 59–68. Elsevier. (also available at https://linkinghub.elsevier.com/retrieve/pii/B9780123847300002561).

Ross, T., Rasmussen, S., Fazil, A., Paoli, G. & Sumner, J. 2009. Quantitative risk assessment of *Listeria monocytogenes* in ready-to-eat meats in Australia. *International Journal of Food Microbiology*, 131(2): 128–137. https://doi.org/10.1016/j.ijfoodmicro.2009.02.007

Ross, T. & Sumner, J. 2002. A simple, spreadsheet-based, food safety risk assessment tool. *International Journal of Food Microbiology*, 77(1): 39–53. https://doi.org/10.1016/S0168-1605(02)00061-2

van Rossum, C.T.M., Fransen, H.P., Verkaik-Kloosterman, J., Buurma-Rethans, E.J.M. & Ocké, M.C. 2011. Dutch National Food Consumption Survey 2007-2010 : Diet of children and adults aged 7 to 69 years. (also available at https://rivm.openrepository.com/rivm/handle/10029/261553).

Rothstein, H.R., Sutton, A.J. & Borenstein, M. 2006. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. H.R. Rothstein, A.J. Sutton & M. Borenstein, eds. Chichester, UK, John Wiley & Sons, Ltd. (also available at http://doi.wiley.com/10.1002/0470870168).

Roulo, R.M., Fishburn, J.D., Amosu, M., Etchison, A.R. & Smith, M.A. 2014. Dose Response of *Listeria monocytogenes* Invasion, Fetal Morbidity, and Fetal Mortality after Oral Challenge in Pregnant and Nonpregnant Mongolian Gerbils. *Infection and Immunity*, 82(11): 4834–4841. https://doi.org/10.1128/IAI.01514-14

Ruzante, J.M., Whiting, R.C., Dennis, S.B. & Buchanan, R.L. 2013. Microbial Risk Assessment. *In* M.P. Doyle & R.L. Buchanan, eds. *Food Microbiology*. Fourth edition, pp. 1023–1037. (also available at http://www.asmscience.org/content/book/10.1128/9781555818463.chap41).
Ryan, M.O., Gurian, P.L., Haas, C.N., Rose, J.B. & Duzinski, P.J. 2013. Acceptable microbial risk: Cost-benefit analysis of a boil water order for *Cryptosporidium*. *Journal - American Water Works Association*, 105(4): E189–E194. https://doi.org/10.5942/jawwa.2013.105.0020
Saltelli, A., Chan, K. & Scott, E.M. 2008. *Sensitivity Analysis*. Second edition. Wiley. 494 pp. (also available at https://www.wiley.com/en-au/Sensitivity+Analysis-p-9780470743829).
Samapundo, S., Climat, R., Xhaferi, R. & Devlieghere, F. 2015. Food safety knowledge, attitudes and practices of street food vendors and consumers in Port-au-Prince, Haiti. *Food Control*, 50: 457–466. https://doi.org/10.1016/j.foodcont.2014.09.010
Scallan, E., Hoekstra, R.M., Angulo, F.J., Tauxe, R.V., Widdowson, M.-A., Roy, S.L., Jones, J.L. & Griffin, P.M. 2011. Foodborne Illness aquired in the United States—Major Pathogens. *Emerging Infectious Diseases*, 17(1): 7–15.
Scallan, E., Hoekstra, R.M., Mahon, B.E., Jones, T.F. & Griffin, P.M. 2015. An assessment of the human health impact of seven leading foodborne pathogens in the United States using disability adjusted life years. *Epidemiology & Infection*, 143(13): 2795–2804. https://doi.org/10.1017/S0950268814003185
Schaffner, D.W. 2003. Models: what comes after the next generation. *In* R.J. McKellar & X. Lu, eds. *Modeling Microbial Responses in Food*. First edition, p. Contemporary Food Science. CRC Press. (also available at https://www.crcpress.com/Modeling-Microbial-Responses-in-Food/McKellar-Lu/p/book/9780849312373).
Schaffner, D.W. 2004. Mathematical Frameworks for Modeling *Listeria* Cross-contamination in Food. *Journal of Food Science*, 69(6): R155–R159. https://doi.org/10.1111/j.1365-2621.2004.tb11004.x
Shinagawa, K., Hu, D. & Yoshida, S. 1997. Correspondence and problem for enterohemorrhagic *E. coli* O157 outbreak in Morioka city, Iwate. *Koshu Eisei Kenkyu*, 46: 104–112.
Shorten, P.R., Pleasants, A.B. & Soboleva, T.K. 2006. Estimation of microbial growth using population measurements subject to a detection limit. *International Journal of Food Microbiology*, 108: 369–375.
Skjerve, E. 1999. Possible Increase of Human *Taenia saginata* Infections through Import of Beef to Norway from a High Prevalence Area. *Journal of Food Protection*, 62(11): 1314–1319. https://doi.org/10.4315/0362-028X-62.11.1314
Slikker Jr., W. 2018. Biomarkers and their impact on precision medicine. *Experimental Biology and Medicine*, 243(3): 211–212. https://doi.org/10.1177/1535370217733426
Smid, J.H., de Jonge, R., Havelaar, A.H. & Pielaat, A. 2013. Variability and Uncertainty Analysis of the Cross-Contamination Ratios of \emphSalmonella During Pork Cutting. *Risk Analysis*, 33(6): 1100–1115.
Smith, A.M., Tau, N.P., Smouse, S.L., Allam, M., Ismail, A., Ramalwa, N.R., Disenyeng, B., Ngomane, M. & Thomas, J. 2019. Outbreak of *Listeria monocytogenes* in South Africa, 2017–2018: Laboratory Activities and Experiences Associated with Whole-Genome Sequencing Analysis of Isolates. *Foodborne Pathogens and Disease*. https://doi.org/10.1089/fpd.2018.2586
Smith, B.A., Fazil, A. & Lammerding, A.M. 2013. A risk assessment model for *Escherichia coli* O157:H7 in ground beef and beef cuts in Canada: Evaluating the effects of interventions. *Food Control*, 29(2): 364–381.
Smith, M.A., Takeuchi, K., Anderson, G., Ware, G.O., McClure, H.M., Raybourne, R.B., Mytle, N. & Doyle, M.P. 2008. Dose-Response Model for *Listeria monocytogenes*-Induced Stillbirths in Nonhuman Primates. *Infection and Immunity*, 76(2): 726–731. https://doi.org/10.1128/IAI.01366-06
Smith, M.A., Takeuchi, K., Brackett, R.E., McClure, H.M., Raybourne, R.B., Williams, K.M., Babu, U.S., Ware, G.O., Broderson, J.R. & Doyle, M.P. 2003. Nonhuman Primate Model for *Listeria monocytogenes*-Induced Stillbirths. *Infection and Immunity*, 71(3): 1574–1579. https://doi.org/10.1128/IAI.71.3.1574-1579.2003

Ssemanda, J.N. 2018. *Towards microbial safety of fresh vegetables in Rwanda*. The Netherlands, Wageningen University. (PhD Thesis). (also available at https://www.wur.nl/en/activity/Towards-microbial-safety-of-fresh-vegetables-in-Rwanda-1.htm).

Ssemanda, J.N., Reij, M.W., Bagabe, M.C., Muvunyi, C.M., Nyamusore, J., Joosten, H. & Zwietering, M.H. 2018. Estimates of the burden of illnesses related to foodborne pathogens as from the syndromic surveillance data of 2013 in Rwanda. *Microbial Risk Analysis*, 9: 55–63. https://doi.org/10.1016/j.mran.2018.02.002

Stephens, P.J., Joynson, J.A., Davies, K.W., Holbrook, R., Lappin-Scott, H.M. & Humphrey, T.J. 1997. The use of an automated growth analyser to measure recovery times of single heat-injured *Salmonella* cells. *Journal of Applied Microbiology*, 83(4): 445–455. https://doi.org/10.1046/j.1365-2672.1997.00255.x

Steyerberg, E.W., Vickers, A.J., Cook, N.R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M.J. & Kattan, M.W. 2010. Assessing the Performance of Prediction Models. *Epidemiology*, 21(1): 128–138. https://doi.org/10.1097/ede.0b013e3181c30fb2

Strachan, N.J.C., Doyle, M.P., Kasuga, F., Rotariu, O. & Ogden, I.D. 2005. Dose response modelling of \emphEscherichia coli O157 incorporating data from foodborne and environmental outbreaks. *International Journal of Food Microbiology*, 103: 35–47.

Sumner, J. 2016. The impact of transport to Australia's distant markets on the shelf-life of beef and sheep primals. No. 2016.1075. North Sydney, Australian Meat Processor Corporation. (also available at https://www.ampc.com.au/uploads/cgblog/id10/2016-1075-The-impact-of-transport-to-Australias-distant-markets-Final-Report.pdf).

Sumner, J. & Ross, T. 2002. A semi-quantitative seafood safety risk assessment. *International Journal of Food Microbiology*, 77(1): 55–59.

Sumner, J., Ross, T. & Ababouch, L. 2004. Application of Risk Assessment in the Fish Industry. Fisheries Technical Paper No. 442. Rome, Italy, FAO [Food and Agricultural Organization of the United Nations].

Tenenhaus-Aziza, F. & Ellouze, M. 2015. Software for predictive microbiology and risk assessment: A description and comparison of tools presented at the ICPMF8 Software Fair. *Food Microbiology*, 45: 290–299. https://doi.org/10.1016/j.fm.2014.06.026

Tessema, A.G., Gelaye, K.A. & Chercos, D.H. 2014. Factors affecting food handling Practices among food handlers of Dangila town food and drink establishments, North West Ethiopia. *BMC Public Health*, 14(1): 571. https://doi.org/10.1186/1471-2458-14-571

Teunis, P.F.M. 1997. Infectious gastro-enteritis - opportunities for dose response modelling. Bilthoven, RIVM. (also available at https://www.rivm.nl/bibliotheek/rapporten/284550003.html).

Teunis, P.F.M., Chappell, C.L. & Okhuysen, P.C. 2002a. Cryptosporidium Dose Response Studies: Variation Between Isolates. *Risk Analysis*, 22(1): 175–185. https://doi.org/10.1111/0272-4332.00014

Teunis, P.F.M., Chappell, C.L. & Okhuysen, P.C. 2002b. Cryptosporidium Dose-Response Studies: Variation Between Hosts. *Risk Analysis*, 22(3): 475–485. https://doi.org/10.1111/0272-4332.00046

Teunis, P.F.M. & Havelaar, A.H. 2000. The Beta Poisson Dose-Response Model Is Not a Single-Hit Model. *Risk Analysis*, 20(4): 513–520.

Teunis, P.F.M., van der Heijden, O.G., van der Giessen, J.W.B. & Havelaar, A.H. 1996. The dose-response relation in human volunteers for gastro-intestinal pathogens. Bilthoven, RIVM. (also available at https://www.rivm.nl/bibliotheek/rapporten/284550002.html).

Teunis, P.F.M., Kasuga, F., Fazil, A., Ogden, I.D., Rotariu, O. & Strachan, N.J.C. 2010. Dose–response modeling of *Salmonella* using outbreak data. *International Journal of Food Microbiology*, 144(2): 243–249. https://doi.org/10.1016/j.ijfoodmicro.2010.09.026

Teunis, P.F.M., Moe, C.L., Liu, P., Miller, S.E., Lindesmith, L., Baric, R.S., Le Pendu, J. & Calderon, R.L. 2008. Norwalk virus: How infectious is it? *Journal of Medical Virology*, 80(8): 1468–1476.

Teunis, P.F.M., Nagelkerke, N.J.D. & Haas, C.N. 1999. Dose Response Models For Infectious Gastroenteritis. *Risk Analysis*, 19(6): 1251–1260.

Thompson, K.M. & Graham, J.D. 1996. Going beyond the single number: Using probabilistic risk assessment to improve risk management. *Human and Ecological Risk Assessment*, 2(4): 1008–1034. https://doi.org/10.1080/10807039609383660

Thomsen, S.T., de Boer, W., Pires, S.M., Devleesschauwer, B., Fagt, S., Andersen, R., Poulsen, M. & van der Voet, H. 2019. A probabilistic approach for risk-benefit assessment of food substitutions: A case study on substituting meat by fish. *Food and Chemical Toxicology*, 126: 79–96. https://doi.org/10.1016/j.fct.2019.02.018

Tomatsis, L., ed. 1990. *Cancer: Causes, occurrence and control*. Lyon, France, International Agency for Research on Cancer. (also available at http://doi.wiley.com/10.1002/sim.4780110216).

USEPA. 2006. Peer Review and Peer Involvement at the U.S. Environmental Protection Agency. United States Environmental Protection Agency. [Cited 12 October 2018]. https://www.epa.gov/sites/production/files/2015-01/documents/peer_review_policy_and_memo.pdf

USFDA. 2005. Quantitative Risk Assessment on the Public Health Impact of Pathogenic *Vibrio parahaemolyticus* in Raw Oysters. USDHHS/FDA [U.S. Department of Health and Human Services / Food and Drug Administration]. (also available at https://www.fda.gov/food/foodscienceresearch/risksafetyassessment/ucm050421.htm).

USNACMCF. 2004. Response to the questions posed by FSIS regarding performance standards with particular reference to raw ground chicken. Atlanta, GA, USA, US National Advisory Committee on Microbiological Criteria for Foods. (also available at https://www.fsis.usda.gov/wps/wcm/connect/e2ac5b0d-0c6a-4075-adb5-eb1ccbc1b1ff/NACMCF_Ground_Chicken_082704.pdf?MOD=AJPERES).

Vally, H., Glass, K., Ford, L., Hall, G., Kirk, M.D., Shadbolt, C., Veitch, M., Fullerton, K.E., Musto, J. & Becker, N. 2014. Proportion of Illness Acquired by Foodborne Transmission for Nine Enteric Pathogens in Australia: An Expert Elicitation. *Foodborne Pathogens and Disease*, 11(9): 727–733. https://doi.org/10.1089/fpd.2014.1746

Van Abel, N., Schoen, M.E., Kissel, J.C. & Meschke, J.S. 2017. Comparison of Risk Predicted by Multiple Norovirus Dose–Response Models and Implications for Quantitative Microbial Risk Assessment. *Risk Analysis*, 37(2): 245–264. https://doi.org/10.1111/risa.12616

Van der Fels-Klerx, H.J., Van Asselt, E.D., Raley, M., Poulsen, M., Korsgaard, H., Bredsdorff, L., Nauta, M., D'agostino, M., Coles, D., Marvin, H.J.P. & Frewer, L.J. 2018. Critical review of methods for risk ranking of food-related hazards, based on risks for human health. *Critical Reviews in Food Science and Nutrition*, 58(2): 178–193. https://doi.org/10.1080/10408398.2016.1141165

Vatanpour, S., Hrudey, S. & Dinu, I. 2015. Can Public Health Risk Assessment Using Risk Matrices Be Misleading? *International Journal of Environmental Research and Public Health*, 12(8): 9575–9588. https://doi.org/10.3390/ijerph120809575

Vimont, A.J.M., Kiermeier, A., Padula, D.J., Holds, G.L. & Pointon, A.M. 2005. The adequacy of sample type/weight and incubation period on detection of *Salmonella* spp. in slaughter cattle. *Journal of Veterinary Diagnostic Investigation*, 17(5): 430–435.

Vose, D. 2008. *Risk Analysis: A Quantitative Guide, 3rd Edition*. Third edition. Wiley. 752 pp. (also available at https://www.wiley.com/en-au/Risk+Analysis%3A+A+Quantitative+Guide%2C+3rd+Edition-p-9780470512845).

Whittemore, A.S. 1983. Facts and Values in Risk Analysis for Environmental Toxicants. *Risk Analysis*, 3(1): 23–33. https://doi.org/10.1111/j.1539-6924.1983.tb00103.x

WHO. 1999. Health-based monitoring of recreational waters: the feasibility of a new approach (the "Annapolis Protocol"). No. WHO/SDE/WSH/99.1. Geneva, Switzerland, World Health Organization. (also available at http://www.who.int/water_sanitation_health/bathing/annapolis.pdf).

WHO. 2000. The world health report 2000 - Health systems: improving performance. Geneva, WHO. (also available at http://www.who.int/whr/2000/en/).

WHO. 2001. Joint FAO/WHO Expert Consultation on Risk Assessment of Microbiological Hazards in Foods : hazard identification, exposure assessment and hazard characterization of Campylobacter

spp. in broiler chickens and Vibrio spp. in seafood, WHO headquarters, Geneva, Switzerland, 23-27 July 2001. No. WHO/SDE/PHE/FOS/01.4. Geneva, WHO48 pp. (also available at http://apps.who.int/iris/handle/10665/67090).
WHO, ed. 2003. *Guidelines for safe recreational water environments: Volume 1, Coastal and fresh water*. Geneva, World Health Organization. 2 pp.
WHO. 2015. *WHO estimates of the global burden of foodborne diseases: Foodborne Disease Burden Epidemiology Reference Group*. World Health Organization. (also available at http://www.who.int/foodsafety/areas_work/foodborne-diseases/ferg/en/).
WHO/FAO, ed. 2006. *Food safety risk analysis: a guide for national food safety authorities*. FAO food and nutrition paper No. 87. Rome, World Health Organization : Food and Agriculture Organization of the United Nations. 102 pp. (also available at http://www.fao.org/3/a0822e/a0822e.pdf).
WHO/OECD. 2003. *Assessing microbial safety of drinking water: Improving approaches and methods*. A. Dufour, M. Snozzi, W. Koster, J. Bartram, E. Ronchi & L. Fewtrell, eds. World Health Organization/Organisation for Economic Co-operation and Development. (also available at http://www.who.int/water_sanitation_health/publications/9241546301/en/).
Wieland, B., Dhollander, S., Salman, M. & Koenen, F. 2011. Qualitative risk assessment in a data-scarce environment: A model to assess the impact of control measures on spread of African Swine Fever. *Preventive Veterinary Medicine*, 99(1): 4–14. https://doi.org/10.1016/j.prevetmed.2011.01.001
Williams, D., Castleman, J., Lee, C.-C., Mote, B. & Smith, M.A. 2009. Risk of Fetal Mortality After Exposure to *Listeria monocytogenes* Based on Dose-Response Data from Pregnant Guinea Pigs and Primates. *Risk Analysis*, 29(11): 1495–1505. https://doi.org/10.1111/j.1539-6924.2009.01308.x
Williams, D., Irvin, E.A., Chmielewski, R.A., Frank, J.F. & Smith, M.A. 2007. Dose-Response of *Listeria monocytogenes* after Oral Exposure in Pregnant Guinea Pigs. *Journal of Food Protection*, 70(5): 1122–1128. https://doi.org/10.4315/0362-028X-70.5.1122
Williams, M.S. & Ebel, E.D. 2014. Fitting a distribution to censored contamination data using Markov Chain Monte Carlo methods and samples selected with unequal probabilities. *Environmental Science & Technology*, 48(22): 13316–13322. https://doi.org/10.1021/es5035574
Williams, M.S., Ebel, E.D. & Vose, D. 2011a. Methodology for determining the appropriateness of a linear Dose-Response function. *Risk Analysis*, 31(3): 345–350. https://doi.org/10.1111/j.1539-6924.2010.01518.x
Williams, M.S., Ebel, E.D. & Vose, D. 2011b. Framework for microbial food-safety risk assessments amenable to Bayesian modeling. *Risk Analysis*, 31(4): 548–565. https://doi.org/10.1111/j.1539-6924.2010.01532.x
Wilson, P.D.G., Brocklehurst, T.F., Arino, S., Thuault, D., Jakobsen, M., Lange, M., Farkas, J., Wimpenny, J.W.T. & Van Impe, J.F. 2002. Modelling microbial growth in structured foods: towards a unified approach. *International Journal of Food Microbiology*, 73(2): 275–289. https://doi.org/10.1016/S0168-1605(01)00660-2
Woods, C.R. 2013. False-Positive Results for Immunoglobulin M Serologic Results: Explanations and Examples. *Journal of the Pediatric Infectious Diseases Society*, 2(1): 87–90. https://doi.org/10.1093/jpids/pis133
Wooldridge, M. 2008. Qualitative Risk Assessment. *Microbial Risk Analysis of Foods*: 1–28. https://doi.org/10.1128/9781555815752.ch1
WTO. 1998. Australia - Measures Affecting Importation of Salmon - Report of the Panel. , p. 267. No. WT/DS18/R. Brussels, Belgium, World Trade Organisztion. (also available at https://www.wto.org/english/tratop_e/dispu_e/cases_e/ds18_e.htm).
WTO. 2000. Committee on Sanitary and Phytosanitary Measures - Guidelines to further the practical implementation of Article 5.5 [G/SPS/15]. World Trade Organization. [Cited 6 December 2018]. https://docsonline.wto.org/Dol2FE/Pages/FormerScriptedSearch/directdoc.aspx?DDFDocuments/t/G/SPS/15.DOC

Wu, F.M., Doyle, M.P., Beuchat, L.R., Wells, J.G., Mintz, E.D. & Swaminathan, B. 2000. Fate of Shigella sonnei on parsley and methods of disinfection. *Journal of Food Protection*, 63(5): 568–572.
Young, I., Reimer, D., Greig, J., Meldrum, R., Turgeon, P. & Waddell, L. 2017a. Explaining Consumer Safe Food Handling Through Behavior-Change Theories: A Systematic Review. *Foodborne Pathogens and Disease*, 14(11): 609–622. https://doi.org/10.1089/fpd.2017.2288
Young, I., Thaivalappil, A., Reimer, D. & Greig, J. 2017b. Food Safety at Farmers' Markets: A Knowledge Synthesis of Published Research. *Journal of Food Protection*, 80(12): 2033–2047. https://doi.org/10.4315/0362-028X.JFP-17-193
Zhang, D., McQuestin, O.J., Mellefont, L.A. & Ross, T. 2010. The influence of non-lethal temperature on the rate of inactivation of vegetative bacteria in inimical environments may be independent of bacterial species. *Food Microbiology*, 27(4): 453–459. https://doi.org/10.1016/j.fm.2009.12.006
Zilelidou, E.A., Tsourou, V., Poimenidou, S., Loukou, A. & Skandamis, P.N. 2015. Modeling transfer of *Escherichia coli* O157:H7 and *Listeria monocytogenes* during preparation of fresh-cut salads: Impact of cutting and shredding practices. *Food Microbiology*, 45: 254–265. https://doi.org/10.1016/j.fm.2014.06.019
Zoellner, C., Jennings, R., Wiedmann, M. & Ivanek, R. 2019. EnABLe: An agent-based model to understand *Listeria* dynamics in food processing facilities. *Scientific Reports*, 9(1): 1–14. https://doi.org/10.1038/s41598-018-36654-z
Zwietering, M.H., Cuppers, H.G.A.M., de Wit, J.C. & van't Riet, K. 1994. Evaluation of Data Transformations and Validation of a Model for the Effect of Temperature on Bacterial Growth. *Appl. Environ. Microbiol.*, 60(1): 195–203.
Zwietering, M.H. & Hasting, A.P.M. 1997a. Modelling the hygienic processing of foods – a global process overview. *IChemE Transactions Part C: Food Bioproducts Processing*, 75(C3): 159–167.
Zwietering, M.H. & Hasting, A.P.M. 1997b. Modelling the hygienic processing of foods – influence of individual process stages. *IChemE Transactions Part C: Food Bioproducts Processing*, 75(C3): 168–173.
Zwietering, M.H., Jongenburger, I., Rombouts, F.M. & Van 'T Riet, K. 1990. Modeling of the Bacterial Growth Curve. *Applied and Environmental Microbiology*, 56(6): 1875–1881.