5. Quantitative risk characterization

5.1 Introduction

Quantitative risk assessment can be either deterministic (meaning single values like means or percentiles are used to describe model variables) or probabilistic (meaning that probability distributions are used to describe model variables). Most of the literature, guidance and the best-known examples in microbiological risk assessment are probabilistic quantitative risk assessments. This approach offers many distinct advantages over deterministic risk assessment, and these are described at length in Chapter 3 and beyond. Examples of deterministic quantitative risk assessment can be found most readily in the food additive safety assessment (also known as chemical risk assessment) literature. FAO and WHO have produced numerous examples of probabilistic risk assessments, as have numerous food safety authorities around the world.

A numerical scale of measure is generally more informative than a qualitative scale. Consequently, a quantitative risk characterization will address identified risk management questions at a finer level of detail or resolution than a qualitative or semi-quantitative risk assessment, and facilitate a more precise comparison between risks and between risk management options. However, the extra level of detail can be at the expense of a far greater time to completion, a reduction in scope and a greater difficulty in understanding the model. Probabilistic techniques are more complicated and therefore introduce a greater risk of error and of not being well understood by stakeholders. In addition, quantitative risk assessments may rely on subjective quantitative assumptions (WHO/OECD, 2003: 80), and the mathematical precision of these quantitative results can inadvertently over-emphasize the real level of accuracy. This has been recognized for a long time in the risk analysis community. Whittemore (1983) noted: "Quantitative risk analyses produce numbers that, out of context, take on lives of their own, free of qualifiers, caveats and assumptions that created them". With these caveats in mind, all else equal, a good quantitative risk assessment is to be preferred over a qualitative or semi-quantitative risk assessment.

A list of desirable properties of quantitative characterizations is given below, followed by a discussion of issues of inter-individual variability, randomness and uncertainty: three aspects of risk quantification that are described by distributions and thus often get confused. Finally, the integration of outputs from exposure assessment and hazard characterization is discussed, including the integration of uncertainty and variability.

5.2 Quantitative measures

Quantitative measures of risk must combine in some form an expression of the two quantitative components of risk, namely some measure of the probability of the risk occurring; and the size of the impact should that risk occur (Kaplan and Garrick, 1981). In this section various ways of combining probability and impact are discussed, and plots of what these can look like graphically are provided for illustration, together with the effect of including uncertainty.

5.2.1 Measure of probability

Probability measures in microbiological food safety risk analysis must relate to a specified level of exposure which can, for example, be the consumption of a particular quantity of food product, being a consumer for a year in a particular country, or an individual exposure event (which may not be the same as consumption if the exposure is indirect). The probability measures are generally expressed in one of two forms:

- the probability of the risk event occurring with a specified exposure event (e.g. probability of illness from consuming a random egg), or within a period (e.g. the probability of getting ill at least once in a year for a random person consuming eggs); or

- the average number of risk events that may occur within a specified period.

There are advantages and disadvantages in selecting each probability measure. The first of these options underlines the probabilistic content of the risk measure, whilst the second can be misread to make one believe that the risk event will occur deterministically with the specified frequency. At the same time, specifying a risk in a probability term makes it difficult to express the possibility of multiple occurrences of the risk event, which is progressively more important with an increase in the estimated expected frequency. For example, if an outbreak is considered to occur randomly in time with an expectation of one event a year, the probability of occurring at least once in a year is about 63%. By contrast a risk with an expectation of twice a year has a probability of occurring at least once in a year of about 86%. The second risk has twice the frequency of the first, but the probability of occurrence does not obviously reflect that. The choice of probability measure needs to be carefully made to make any explanation of the risk assessment results as clear as possible to the intended audience.

5.2.2 Measure of impact

The selected measure or measures of impact will reflect what the risk manager cares about. It could be a case of human illness or of death, but the cases of illness could be further stratified into various levels of illness if the decision-maker values the distinction. There may also be a translation from an illness into an economic impact, or into some social impact measure, like quality adjusted life years (QALY) discussed further in Section 7.2.1.

5.2.3 Measures of risk

The risk measure combines the probability and impact components discussed above to provide a description of the risk, together with attendant uncertainties. The selected option needs to be chosen to make the risk estimate the most readily understood by the intended audience. It may therefore be very useful to produce more than one expression for different audiences. The choice of expression should also be the most useful for the decision-maker (for example if one is making comparisons with other risks, the measures should be consistent). There are also issues that should be borne in mind in communicating risk estimates to stakeholders regarding how people react to different expressions of the same risk, which can get in the way of constructive dialogue. For example, informing the public of a country of 20 million people that one has estimated that there is a one in a million chance that a person will die from a particular hazard per year may generate an entirely different response from informing them that one has estimated that, on average, 20 people will die a year from the hazard. There is a considerable body of available literature on risk perception and interpretation that risk assessors and risk managers should make themselves familiar with.

The risk measure may be a single point probabilistic measure, as discussed below; for example, the probability of at least one illness within a certain period or the expected number of cases next year. This means that, if no uncertainty has been included in the risk assessment model or if uncertainty and randomness have been combined, these outputs are fixed values (Figure 5.1). If uncertainty has been put into the model and not combined with randomness, the outputs are uncertainty distributions (Figure 5.1a).

The risk measure may alternatively be a probability distribution, for example, a probability distribution of the number of adverse health events a random person might experience next year. This will be a first-order distribution if no uncertainty has been included in the model (Figure 5.1b), or if uncertainty and randomness have been combined. If uncertainty has been put into the model and not combined with randomness, the output will be a second-order probability distribution (Figure 5.1c).

Thirdly, the risk measure may describe the variation in risk across a population. That risk can, for example, be characterized as the probability of illness per serving. We can end up with a distribution of the variability across subpopulations of that probability, perhaps because some subpopulations have a more highly-contaminated source of food, or they prepare or handle the food differently due to custom, or their dose-response curve is steeper than others because they are more susceptible to a bacterium. One can graph the variation in that probability per serving from one subpopulation to another if it is illuminating to compare subpopulations. If the risk assessment did not include uncertainty, we could use a single probability measure to describe the risk for each subpopulation (Figure 5.1d). If the risk assessment has included uncertainty and not combined the uncertainty into the probability measure, we can also look at how sure we are about these estimates of probability per serving (Figure 5.1e). It is difficult to graphically compare two or more second-order distributions so, whilst it is theoretically possible to produce, for example, probability distributions of the number of illnesses a person or subpopulation may endure over a period, if these are second-order distributions it will generally be far clearer to make a comparison of an appropriate statistic (mean, 90th percentile, etc.) with attendant uncertainties.

Risk per serving

The risk per serving suffers the ambiguity of "What should be defined as a serving?" Thus, standard quantities need first to be defined, such as a serving being 100 g of cooked chicken or 150 ml of orange juice, or a probability distribution of serving size. The risk is also not easily translated to any individual, as one needs to take into account the amount of that food that an individual might consume within a defined period. However, if a standard quantity (like 100 g cooked weight, or 30 g protein intake, or 1000 calories) is first set, the risk per serving measure provides an easy comparison of the risk from direct consumption of different food products. It can also be helpful in establishing cost–benefit type arguments where, for example, one is looking for the lowest risk for a given nutrition requirement.

Individual risk

Individual risk can be specified for a random individual within the population of concern, or for a random consumer of the product (assuming not everyone in the population consumes the
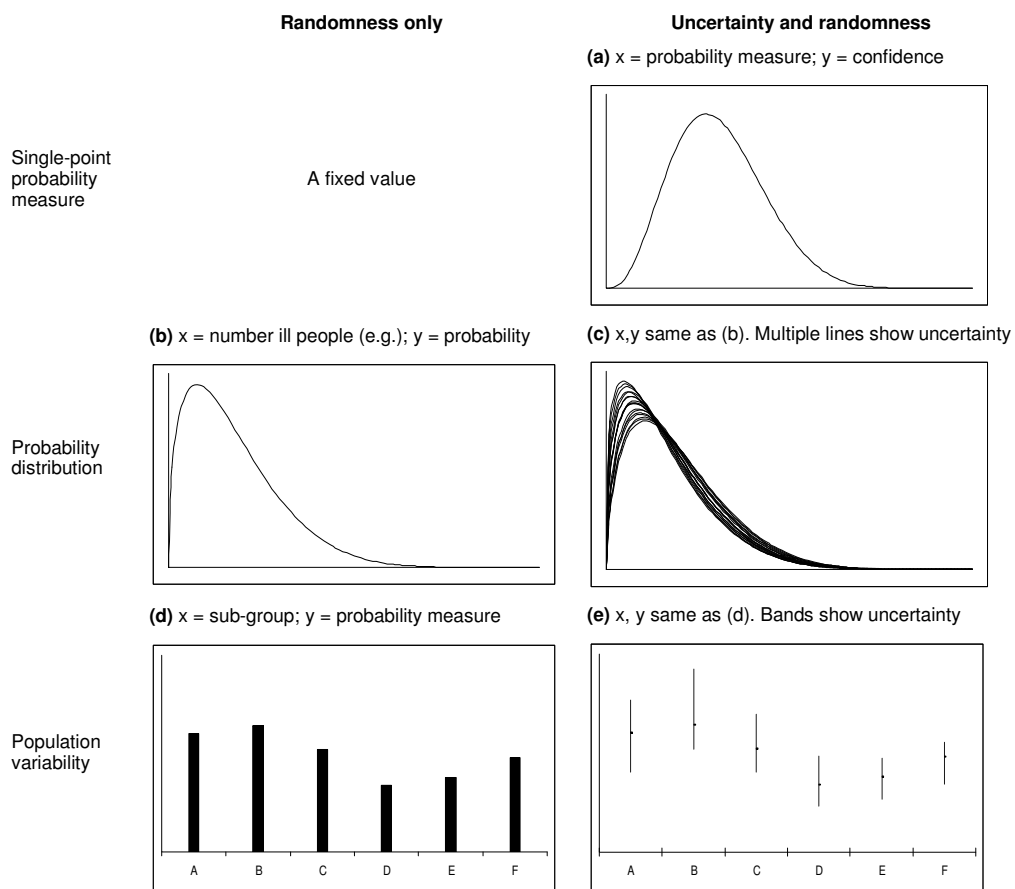
Randomness only   **Uncertainty and randomness**

                                                    (a) x = probability measure; y = confidence

Single-point
probability   A fixed value
measure

(b) x = number ill people (e.g.); y = probability   **(c)** x,y same as (b). Multiple lines show uncertainty

Probability
distribution

(d) x = sub-group; y = probability measure   **(e)** x, y same as (d). Bands show uncertainty

Population
variability

Figure 5.1 A matrix of various types of quantitative outputs one can produce from a risk assessment describing variability, randomness and uncertainty.

product in question, and that only those consuming are at risk, i.e. that there are no significant secondary infections or cross-contamination[1]). It can also be specified for random individual of various subpopulations when one wishes to explore the degree to which subpopulations differ in bearing the population risk. The following are some examples of individual risk estimates:

- The probability per year that a random individual will suffer illness X from exposure to bacteria Y in food Z.

---

[1] Cross-contamination is discussed in the FAO/WHO guidelines on Exposure Assessment of Microbiological Hazards in Foods (FAO/WHO, 2008).

- The probability per year that a random individual will suffer any deterioration in health X from exposure to bacteria Y in food type Z.

- The probability that a person will suffer some adverse health effect in their lifetime from exposure to bacteria Y in foods.

- The expected number of foodborne-related adverse health events for a random individual from consuming food type Z in a year.

- Distribution of the number of foodborne-related adverse health events for a random individual from consuming food type Z in a year.

- Per capita (or per kg consumed, or per kg produced, by the nation) expected incidence of health impact X from food type Z.

Risk per person is very often a very low number (for example 0.000013 expected illnesses per person per year), making it difficult to understand and compare, but the numbers can be raised to more useable figures by considering the risk over a large number of people, for example, by changing the estimate above to 1.3 expected illnesses per 100 000 people per year. The multiplying factor can be chosen to make the risk measure more accessible: for example, 100 000 might be selected because it represents the size of a small city, and thinking of '1.3 illnesses per year for a small city' has a mental picture attached to it that '0.000013 expected illnesses per person per year' does not.

Population-level risk

A population-level risk estimate considers the risk distributed over the population or sub-population of interest, and might also look at the risk burden absorbed by the population as a whole. It does not distinguish between sub-groups within that population, such as by region, ethnicity, age or health status. The following are some examples of population-level risk estimates:

- Total number of cases of foodborne illness one might expect within the population in a year.

- Expected number of hospital bed-days taken up per year as a result of a particular foodborne pathogen.

- The number of QALYs or money lost per year from foodborne pathogen in a particular food type.

- Probability that there will be at least one outbreak (or one death, one illness, etc.) in the population in a year.

- Probability that there will be more than 10 000 illnesses in the population in a year.

These estimates can be produced for separate subpopulations if required, and aggregated to a single measure for the population as a whole.

5.2.4 Matching dose-response endpoints to the risk measure

Exposure to microbiological agents can result in a continuum of responses ranging from asymptomatic carriage to death. Risk characterization needs to consider the measurement endpoint (reported health outcome) used in developing the dose-response relationship, and may require estimating the desired risk assessment endpoint(s) from a more or less severe measurement endpoint. A fraction of exposed individuals will become infected. Infection may

be measured as the multiplication of organisms within the host, followed by excretion, or a rise in serum antibodies. A fraction of those infected will exhibit symptomatic illness (the morbidity ratio), as measured by clinical observation or reported by patients or consumer responses. A fraction of those becoming ill will suffer severe symptoms (e.g. bloody diarrhoea), require medical care or hospitalization, or will die (the mortality ratio). Care must be taken to ensure that the implications of the case definition used in a clinical trial or epidemiological investigation are understood. For clinical trials, typical measurement endpoints include infection (as indicated by a faecal positive) or illness (as indicated by diarrhoea). Epidemiological surveys may provide information on morbidity and mortality ratios. It is conceivable that the ratios might be dose-dependent, however, and epidemiological data will not inform this relationship. In some cases, clinical trials have used a continuous dose-response measurement endpoint (e.g. volume of diarrhoea excreted) that might provide some insight about the dose-dependency of outcome severity.

5.2.5 Accounting for subpopulations

Subpopulations of consumers may vary with respect to susceptibility, exposure, or both. If the risk characterization seeks to distinguish risk by subpopulation (e.g. by age class), the exposure assessment output should be disaggregated to reflect variation in exposure among sub-populations (e.g. the frequency, size and preparation of servings consumed by members of each age class). As discussed earlier in relation to variability and uncertainty, if sufficient information is available to develop subpopulation-specific dose-response relationships (e.g. susceptible versus non-susceptible), then the exposure assessment output for each subpopulation can be propagated through its corresponding dose-response model. However, even in cases where such separate dose-response relationships cannot be specified, it may be informative to characterize risk by subpopulation. For example, there may be sufficient data to develop sub-population-specific morbidity or mortality ratios. It should be noted that subpopulations of particular concern (e.g. susceptible consumers) may not correspond directly to easily identified categories (e.g. age classes). Care must therefore be taken to ensure that there is a reasoned basis for classifying consumers as members of different subpopulations, and that subpopulation definitions are consistent between the exposure and dose-response analyses. An example using subpopulations is discussed below (Section 5.5.7).

5.3 Desirable properties of quantitative risk assessments

Quantitative risk assessment includes identification, selection and development or modification of one or more models that are then combined into a framework. A key consideration in choosing appropriate models is the level of detail required for the assessment, consistent with the assessment objectives.

The choice of quantitative model must evaluate how well the model is supported by the available data, how effective the outputs are in informing decision-makers, and how many assumptions have been made in creating the model and the robustness of those assumptions. Inevitably, the process of choosing models, selecting and analysing data, and applying the data and models to answer assessment questions, involves subjective judgement.

Section 6.3 discusses sensitivity analysis, which helps identify key variables that are potentially controllable and can be used to identify key sources of uncertainty for which additional research or data collection could improve the state of knowledge and thereby reduce ambiguity regarding the characterization of risk and comparison of risk management options.

Based upon the above, the key properties of quantitative risk assessment include:

- Clearly defined objectives of the assessment.

- Well-specified scenarios.

- Appropriately selected models, supported as far as possible by data.

- Level of detail of analysis appropriate to the level of the assessment (e.g. screening versus refined).

- Evaluation of uncertainty in scenarios.

- Evaluation of uncertainty in models.

- Explanation of all assumptions and choice of data used in the analysis.

- Quantification and evaluation of randomness, variability and uncertainty in model predictions.

- Proper integration of exposure assessment and hazard characterization to characterize risk.

- Identification of key opportunities for risk mitigation.

- Identification of key opportunities for reducing uncertainty.

- Identification of appropriate risk metrics.

5.4 Variability, randomness and uncertainty

Quantitative risk assessments aim to predict what will happen in the future, or to predict what the effect might be if we were to change the world in some way. The pathways from micro-organisms growing in a food-producing animal to human exposure to these microorganisms and subsequent health effects involve many random processes. A quantitative model uses probability to describe this randomness. This leads to results such as the probability of a randomly selected individual being infected in a given year, or a probability distribution of the number of illnesses there may be in a future period. There is also a great deal of inter-individual variability between animals, farms, human behaviour, etc. Where this variability influences the evaluation of the risk, a quantitative model describes the variability using frequency distributions. The complexity of the system as well as our imperfect measurement methods also leaves us uncertain about the exact values of parameters that would describe the proposed risk pathways. A quantitative model describes this uncertainty with uncertainty distributions, determined by various statistical methods. There are several texts that deal with modelling uncertainty, variability and randomness. Here, an overview of the key concepts is presented, using illustrative examples where necessary. For more details of the methodology the reader is referred to texts such as Morgan and Henrion (1990), Vose (2000) and ModelAssist (2004).

5.4.1 Modelling variability as randomness

Variability is often confused with randomness. If we have assigned some frequency distribution to describe the inter-individual variability of the food-producing animal (mass of a chicken carcass, for example), then a randomly sampled chicken carcass will have a mass given by this same distribution. The frequency distribution is *re-interpreted* as a probability distribution because of the action of a *random* sample. Thus, within quantitative models, some sources of variability can be treated as random variables, thus allowing random sampling from associated probability distributions. A rough rule of thumb would be that one can model variability as

randomness if the number of randomly sampled individuals is very much smaller than the population. For example, few people a year are exposed to *E. coli* O157:H7, so one could consider a person so exposed to have a susceptibility to the bacterium that is drawn at random from the distribution of variability of susceptibility across the entire population. However, it is not appropriate to do this for some sources of variability, and in this case, stratification of the population must be undertaken, and these strata must either be modelled in parallel to give separate estimates of probability, or be weighted to give one estimate of probability. Examples are: stratifying the population according to susceptibility; and stratifying the food product according to producer.

Variability is important because, for example, it reflects the fact that different individuals are subject to different exposures and risks, and different food handling methods produce different levels of risk. An understanding of inter-individual variability will provide insight regarding subgroups of exposed populations that are most exposed or subject to risk, and methods that are more or less dangerous than average. If there are interventions that can be implemented it may be useful to target specific strata (e.g. children vs. adults). In addition, through implementation of intervention strategies (e.g. practices, technologies) aimed at modifying controllable variation (e.g. reduce occurrence of high values of storage time and/or temperature to reduce growth of pathogens during storage), it could be possible to reduce the highest exposures and therefore to reduce risk.

5.4.2 Separation of variability and randomness from uncertainty

Variability and randomness are real-world phenomena. The degree of variability between individuals, animals, bacteria, processing plants, refrigerators, etc., exists whether we have information about it or not. The same applies for probability: bacterial growth, amount of food ingested at a sitting, whether a food item leaves a slaughter plant contaminated, the number of bacteria at the moment of ingestion, etc., are all random (stochastic) variables, and as such are characterized by probability distributions that exist whether we know them or not. Uncertainty, in contrast, is a subjective quality, in that it is a function of the amount of information available to the assessor. Different assessors with different amounts of information may produce different distributions of uncertainty.

Randomness, variability and uncertainty can all be described by distributions that, in essence, look the same. The difference is that the vertical scales describe different quantities: for inter-individual variability distributions, it is relative frequency; for probability distributions it is probability or probability density; and for uncertainty distributions it is relative confidence. The three distinct uses of distributions can be confusing and may lead to them being modelled together in one Monte-Carlo model. However, the result of such a model will describe a mixture, which can be difficult to interpret. Moreover, failure to maintain separation between variability and randomness (the true world), and uncertainty (our level of knowledge), can profoundly affect the risk characterization in some cases (Nauta, 2000). For these reasons, it is considered useful to separate them as far as possible. This can be achieved in a number of ways including second-order modelling (see Box 5.1).

Data are required to define the distributions associated with the model inputs. The available data may be ambiguous and thus it may be difficult to determine whether variability or randomness, or both, are described within it. In subjective estimation of random quantities, it is also usually very difficult to separate the random and uncertainty components in the estimate. Thus, although useful, separation of uncertainty from the other model components may be difficult and is only necessary if the decision is affected. There may be a tendency to ignore uncertainty when producing a probability model, especially if one is not intending to make a second-order model, but uncertainty should not be excluded unless an analysis shows its exclusion to have minimal

> **Box 5.1 Second-order models**
>
> Probability distributions are derived from data. These data are likely to be a sample of some kind and thus when we derive the associated probability distribution, there will be some level of uncertainty associated with it. Overlaying this uncertainty results in a second-order distribution. Visually, a second-order distribution is shown by multiple probability curves on the same graph. Each curve represents a probability distribution and the difference between the curves shows the uncertainty associated with this randomness.

impact, as this exclusion could lead to over-confidence in the accuracy of the model outputs. Examples of recent microbiological risk assessments where separation has been considered include Nauta et al. (2001), Hartnett (2001) and US FDA-CVM (2000).

5.5 Integration of hazard characterization and exposure assessment

Codex guidelines describe the need to assess exposure to a pathogen and assess the level of risk (dose-response relationship) that that exposure represents. Most quantitative risk assessments will have implemented the exposure and dose-response models separately, and risk characterization requires that these are connected together to estimate the risk. In doing so it is crucial that the dose concepts applied in both are mutually consistent with respect to the units of dose and any probability assumptions. This consistency should be included in the planning stage of the modelling whenever possible, to avoid having to adjust the output of exposure or the input of the hazard characterization to achieve consistency, which may not work.

When there is a logical separation between variability and uncertainty in either the exposure assessment or hazard characterization, this distinction should be propagated through the process of integration to determine both the variability and uncertainty in the relevant measures of risks that are the focus of the assessment. Failure to maintain separation between variability and uncertainty can profoundly affect the risk characterization (Nauta, 2000). Additionally, assumptions implicit to specific dose-response models or potential biases associated with estimation of the dose-response can limit the manner in which exposure and dose-response can be combined. Attention to modelling assumptions and potential biases of the dose-response are necessary to ensure a logical integration of exposure and hazard characterization.

In this section the dose concepts as formulated in the FAO/WHO guidelines on exposure assessment and hazard characterization (FAO/WHO, 2003, 2008) are reviewed and suggestions are offered to address the issues of maintaining consistency of units, dose-response model rationales and reducing biases when integrating potentially inconsistent exposure and hazard characterizations.

5.5.1 Units of dose in exposure assessment

According to Codex (CAC, 1999) the output of the exposure assessment is defined as an estimate, with associated uncertainty, of the likelihood and level of a pathogen in a specified

consumer portion of food. With respect to pathogens occurring at relatively low concentrations, this exposure estimate is commonly represented by a prevalence representing the probability that a randomly selected portion of food is contaminated with the pathogen, combined with a probability distribution representing the numbers (or concentration) of pathogens in those portions of food that are contaminated (i.e. contain one or more cells of the pathogen). It is desirable that both the prevalence and the conditioned probability distribution of contamination be presented with attendant uncertainty (FAO/WHO, 2008). For pathogens occurring at relatively high concentrations, the prevalence in consumer portions may be virtually 100%. In that case the determinant aspect of exposure is just the estimated distribution of microbiological concentrations over all consumer food portions.

Whether the level of contamination is expressed as a concentration (CFU per gram or per ml) or a number (CFU) is important when linking this exposure output to a dose-response model. Numbers of CFU potentially ingested are necessarily positive integers. Consequently, a discrete distribution is the most natural choice for the estimated exposure. The use of a continuous distribution for modelling of individual exposures would be most appropriate when pathogen concentrations are relatively high, but can always be converted back to a discrete distribution with some rounding function. Continuous distributions are often used for bacterial counts because they are a lot more flexible and easier to manipulate than discrete distributions. If a concentration is used to express the level of exposure, the concentration has to be multiplied by the amount of food ingested to determine the individual exposure. If the concentration being modelled is in the form of a probabilistic mean, then one needs to use dose-response functions for which inputs are probabilistic (usually, Poisson) mean doses rather than dose-response functions whose input is an actual dose.

5.5.2 Units of dose and response in dose-response assessment

Typically dose-response models in microbiological risk assessment apply the concepts of non-threshold mechanisms, independent action and the particulate nature of the inoculum (FAO/WHO, 2003). This results in the application of single-hit models like the exponential model, the Beta-Poisson model, the Weibull-Gamma model and the hypergeometric model (Haas, 1983; Teunis and Havelaar, 2000). These models assume each ingested cell acts independently, and all cells have the same probability of causing infection. The non-threshold assumption implies the existence of some level of risk for any dose greater than zero.

The FAO/WHO guidelines on hazard characterization (FAO/WHO, 2003) provide a review of current dose-response models. The two principle types of data useful for developing a dose-response assessment are clinical feeding trial studies with human volunteers and epidemiological data on disease incidence associated with foodborne exposure. These different types of human data have varying strengths and weaknesses.

When available, clinical feeding trials data typically consist of measures of illness outcome in small samples of young healthy volunteers exposed to varying levels of a surrogate pathogen(s). Typically, such studies are conducted with stomach acid neutralization by co-administration of an antacid (e.g. bicarbonate) to enhance the survival of the organism in the gastrointestinal tract and minimize inter-individual variation of the 'effective' exposure. Consequently a dose-response relationship estimated on the basis of such data is potentially biased relative to the dose-response associated with foodborne exposures in a population with susceptible as well as healthy individuals. The limited number of participants in feeding trials also means that it has not been practical to observe low rates of infection for low doses and these probabilities have to be extrapolated from higher doses. For bacterial pathogens,

individual exposures within the same dose group of a trial are variable due to randomness of the inocula within the delivery media, which is accounted for in the analysis but adds an extra layer of uncertainty. For some other types of pathogens, such as the protozoa *Cryptosporidium parvum*, the number of organisms can be counted directly and there may be no uncertainty with respect to actual individual level exposures. The functional form of the fitted dose-response model must be aligned with the form of the experiment: for example, a Beta-Poisson dose-response model assumes that the actual dose is a Poisson random variable with known expected value, which is an appropriate model to use for bacterial feeding trials; and the Beta-Binomial dose-response model assumes that one knows the exact number of pathogenic organisms ingested, which is suitable for a feeding trial where the dose has been counted.

Epidemiological data typically consists of a collection of culture-confirmed or otherwise identified illnesses recorded over a specific period and geographical region by public health authorities. Such data may be the result of active or passive ongoing surveillance or specific outbreak investigations. Depending on the pathogen, only a fraction of the identified illness burden may be attributable to foodborne exposures. Additional information is needed to inform hazard characterization to estimate the number of exposures corresponding to any given number of confirmed illnesses, and the likely levels of exposure that occurred. Furthermore, given varying severities of illness that may occur, the number of identified illnesses is only a subset of all illness. The proportion of total illnesses that is subsequently culture-confirmed or otherwise identified is likely to vary substantially for different pathogens as a consequence of differences in virulence and/or host susceptibility (Mead et al. 1999). An advantage of epidemiological data is that one considers the exposure of people who would never be involved in feeding trial experiments: pregnant women, old and infirm, young children, etc.

Data obtained from animal studies are also of value, but a dose-response relationship based on such data is more problematic than would be the case with either experimental or observational data in human populations. However, when experimental human feeding trials data are lacking or epidemiological data are limited and insufficient to determine a dose-response, then the hazard characterization may be largely based on extrapolation from animal studies. In such cases the associated uncertainties of the dose-response assessment are substantially greater, and particular attention should be given to appropriately assess and propagate these uncertainties through the risk characterization. That said, it is considerably more difficult to assess the uncertainty associated with a species-to-species extrapolation than between a small sample of human data extrapolated to a population.

5.5.3 Combining Exposure and Dose-response assessments

An important concern when combining exposure and dose-response assessments is maintaining consistency. First of all, exposure assessment and hazard characterization should deal with, or be applicable to, the same hazard, the same population or subpopulations and the same time frame. This may seem obvious, but due to a lack of data one might choose, for example, to use a surrogate microorganism for the dose-response, or to extrapolate a dose-response relationship estimated based on data obtained with young healthy volunteers to a less homogenous population that includes susceptible individuals. Such extrapolations should be avoided, if at all possible, by looking at alternative modelling approaches, but if this is done it should be clearly stated, and if possible the potential biases and uncertainties of such extrapolation should be incorporated as part of the assessment.

The appropriate combination of the two assessments may depend on whether the dose-response is inferred from individual (feeding trials) or aggregate (epidemiological) level data.

The output of the exposure assessment should be in units of ingested organisms (CFU, cells, PFU [plaque-forming units, used to quantify virus concentrations]) per individual and usually on a per-exposure event basis due to the acute nature of microbiological risks. In contrast, the input of the dose-response may not be on a per-individual level. For example, the exposure may be expressed as a mean or other summary of a distribution of exposures over a group of individuals, though this should be avoided if at all possible. Differences between individual- and group-level exposure summaries in a hazard characterization may create problems of consistency when combining the two assessments for the purpose of risk characterization.

Technically, exposure assessment and hazard characterization can be combined in a Monte Carlo simulation by calculating a probability of infection (or illness) associated with each of k samples from the exposure distribution. For a given sample of n_i cells from the exposure distribution, P(infection| n_i), the probability of infection conditional on the specified dose, would then be calculated based on the estimated or inferred dose-response relationship. The unconditional probability of infection given exposure would then be calculated by taking the mean of the k values of P(infection|n_i) sampled in the Monte Carlo simulation. In such calculations exposure and risk predictions will generally be uncertain due to the epistemic uncertainty associated with alternative models of the exposure distribution and the risk of infection (or illness) at any specified dose level. These uncertainties extend to predictions of risk when the exposure and dose-response are combined and should be properly represented in the output of the assessment.

5.5.4 Dose-response model assumptions

Many of the most common single-hit dose-response models (e.g. the Beta-Poisson and Exponential) assume a Poisson distributed dose to derive a relationship between mean dose and probability of an adverse health effect. This Poisson distribution of dose will not usually be compatible with the distribution of dose that results from the exposure assessment: the most common exception being a dose from a homogeneous food like a liquid or ground meat, where the pathogenic organism may be randomly distributed in the food without clustering. The Poisson-based dose-response model is appropriate to use in the statistical analysis of feeding trial data where the administered dose is a sample from a solution of particular bacterial concentration and the dose may thus be considered Poisson distributed. The parameters of the dose-response function estimated in the statistical analysis can still be used in another (the beta-binomial) dose-response model that assumes the dose is exactly known.

The equation that is the basis of all single-hit models is an expression for the probability of one or more hits occurring under the assumption of independent action. Under this assumption, the probability of infection is expressed by the binomial dose-response model:

$$P_{inf} = 1 - (1-p_m)^n$$

where P_{inf} is the probability of infection (or a more severe health effect), p_m is the probability that a single cell causes infection, and n is the number of pathogens ingested (FAO/WHO, 2003. 2008).

If the host-to-host variability of probability of infection by a single cell is expressed by a Beta distribution, p_m can be replaced by a Beta distribution to incorporate the effects of inter-individual variability. The resulting dose-response relation is a Beta-Binomial model.

If the dose n is not known but assumed to be Poisson distributed with known mean, and p_m is considered constant, the dose-response model is the Exponential.

If the dose is known but dose is assumed to be Poisson distributed with known mean, and p_m is considered to vary with a Beta distribution, the dose-response model is beta-Poisson which would be most appropriately obtained from a fit of the confluent hypergeometric function to the available data (Teunis and Havelaar, 2000).

Effects of uncertainty of the fitted parameters of the dose-response equation should be propagated through the risk characterization calculations.

If the exposure distribution is Poisson, or the coefficient of variation (i.e. the standard deviation divided by the mean) is small, it is sufficient to use the mean level of exposure as an input to the dose-response relation. If the coefficient of variation is large, then the whole exposure distribution has to be taken into account and to ensure consistency, attention should be given to the rationale underlying the chosen dose-response model(s).

5.5.5 Exposure expressed as prevalence

If the quantitative levels of exposure are not known, and exposure is expressed only as a prevalence (for example, of positively tested food items), then dose-response models relating quantitative levels of exposure to the probability of effect cannot be applied. The same situation applies if a dose-response relation as described in the FAO/WHO guidelines on hazard characterization is lacking.

In these situations one might need to utilize data available from monitoring or surveillance, or both, on the prevalence of exposure, and relate it to the incidence rate of illness. Preferably, this prevalence should be the prevalence of a pathogen as close to consumption as possible but where the measured individuals are still representative of the food source as a whole and where the food is the primary vehicle for human exposure to the pathogen.

In some cases, the relationship between this prevalence and the expected rate of illness associated with the food source can be assumed to be linear. This assumption should preferably be supported by the data and the logistics of food processing and handling: cross-contamination between and mixing of food units after the point of prevalence measurement should be negligible.

Under these assumptions, one may be able to establish one of two relationships between prevalence and rate of illness, as presented in Figure 5.2. If a risk characterization is being conducted for a pathogen and food item that is considered the only vehicle for the pathogen (as might, for example, be applicable to *Salmonella* Enteritidis in eggs), then one might expect, to a first approximation, a linear relationship that goes through the origin, where zero prevalence predicts zero illnesses. If the food item under consideration is only one of many routes of exposure, then the rate of illness will not necessarily be zero when the prevalence of the pathogen in the food item is zero, resulting in the upper line in Figure 5.2.

If these assumptions are reasonable, this type of relationship may help predict the effect of lowering the prevalence through intervention in the food chain at a point before the prevalence is measured. An additional concern, however, is that an intervention that lowers the level of pathogens in the food item need not have an effect on the prevalence, but would have an effect on the dose distribution in exposures, and therefore on the human health effect. For this reason, a risk characterization based on a prevalence dose-response relationship must be interpreted with caution.
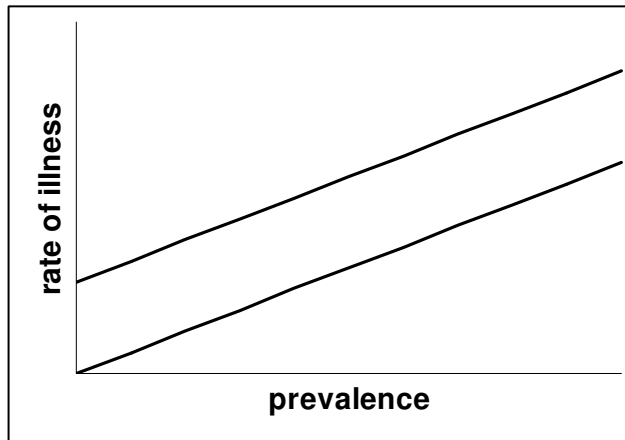
Figure 5.2 Linear prevalence mean risk dose-response relationships, where the lower line represents a single route of exposure, such that zero prevalence means zero risk, while the upper line means multiple routes of exposure, such that zero prevalence in one food does not mean no illness in the population.

The sensitivity and specificity of the test used to measure the prevalence is an issue of concern when using data on prevalence for any aspect of risk characterization and, for that matter, the evaluation of bacterial levels in food. In comparing prevalence, the sensitivities of the methods used should be identical and, if not, they should be known, to allow a correction to the 'true prevalence'. An underlying concern here is that of the microbiological limit of detection (LOD) of the pathogen in the food. If this detection limit varies among methods, then this is likely to have an effect on the prevalence measured. Statistical methods are available to correct for these measurement errors (Gibbons and Coleman, 2001).

If it is known from the exposure assessment that the level of exposure is low, and the only dose-response relationship available is of the type described here, one has to be particularly careful of the effect that the detection limit and sensitivity of the test may have on the prevalence measured to obtain the data used for the dose-response relationship. Low levels of 1 to 10 CFU per 100 g are seldom measured, but may have an impact on the rate of illness of pathogenic microorganisms. Therefore it is generally not advised to combine low level exposure assessments with a dose-response relation of the type discussed here, unless it can be assumed that the distribution of pathogen in an exposure event remains constant for all risk management strategies, or that the exposure level will remain within a straight-line section of the dose-response curve.

5.5.6 Epidemiological-based dose-response relationships

Since aggregate-level (epidemiological) data typically relate observed or inferred mean risk to observed or inferred mean exposure, estimating the relationship between exposure and risk at the individual level from such data can be problematical. This is because the units of measurement (aggregates in the population) are not the same as the targets of inference (individuals). In the epidemiology literature, this inference problem is generally referred to as the problem of ecological, aggregation or cross-level bias (Piantadosi, Byar and Green, 1988;

Greenland and Morgenstern, 1989; Richardson, Stucker and Hemon, 1987). Ecological bias includes the potential effect of confounders, but it is recognized that aggregation itself, in absence of any confounders, can bias aggregate-level relationships relative to underlying individual-level relationships. This is generally referred to as aggregation or cross-level bias.

There is no generally accepted 'solution' to this problem of bias in interpretation of aggregate-level data. However, choosing an appropriate definition of 'mean' risk or exposure, or both, for the groupings (e.g. geometric versus arithmetic mean) of the data used in hazard characterization may help to reduce the effects of bias to a practical level when the derived relationship is intended to represent risk versus dose at the individual level (Haas, 1996; Crump 1998; Guo et al., 1998). Alternative and more sophisticated approaches to reducing the effects of ecological bias exist (King, 1997) but such methods may not be applicable to microbiological risk assessment given the nature and extent of other biases that may be present.

The effects of cross-level bias are an issue of potential concern regardless of the form of the exposure assessment output, be it low-level, high-level or prevalence-based exposure. However, given the nature and extent of other uncertainties, the effect of cross-level bias may be particularly relevant when the dose response is integrated with an exposure assessment where exposures are quantitative at a high level.

5.5.7 Integration of variability and uncertainty

The way in which results from exposure assessment and hazard characterization are integrated will depend on the approach that has been taken to account for variability and uncertainty. The approach taken at each stage should be consistent so that, for example, if exposure has been stratified according to the susceptibility of different populations, there should be a dose-response model for each population. It is important to correctly match these model characteristics when combining results. For example, if variability between subpopulations has been accounted for, the probability distribution for exposure in each individual population should be propagated through the dose-response model for that population.

These ideas are illustrated in Figure 5.3. 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 that it is important to link the correct exposure model with the correct dose-response model if exposure and dose-response are stratified in this way.

The ideas of linkage are further illustrated in Figures 5.4 to 5.7, taking account of variability and uncertainty. In particular, they show results from example models in which stratification of the population and uncertainty have been incorporated to different degrees. In each case, the exposure assessment yields the probability that a randomly selected serving of the food product is contaminated with the microorganism and a probability distribution for the log number of organisms in such a serving. It has been assumed that any variability in contamination across, for example, seasons, areas of the country or food producers has been accounted for via averaging, and thus stratification is not required. The hazard characterization gives the dose-response model. It should be noted that the graphs showing the dose-response models are not probability or uncertainty distributions: they are mathematical functions. Finally, the risk characterization presents two measures of risk. The first measure is at the individual level and is given by the probability that a randomly selected individual becomes ill from consuming a serving of the food product. The second measure is at the population level, defined as the number of cases in the next year.

Uncertainty is not included in the models shown in Figures 5.4 and 5.5. This results in point values for the probability of contamination and individual risk, single-dose-response models and single-probability distributions for the log number of organisms and population-level risk. If, for example, large samples were available, and thus randomness and differences between populations dominant, this approach would be appropriate. Alternatively, results like these would be obtained if the effect of uncertainty is to be investigated by changing the model parameters.

Second-order models are represented by Figures 5.6 and 5.7. This means that there is an uncertainty distribution for both the probability that a serving is contaminated and the individual risk (note the y-axis on each of these graphs – it shows confidence rather than probability). There is also uncertainty associated with the dose-response model and the probability distributions for the log numbers per serving and the population risk. This uncertainty is indicated by the multiple curves on each graph. This approach is appropriate if the uncertainty is large and can be explicitly separated from variability at all stages.

The difference in results when stratification of the population is included compared with not included can be seen by comparing Figures 5.4 and 5.5 (without uncertainty) and Figures 5.6 and 5.7 (with uncertainty). For the models in Figures 5.5 and 5.7, it has been assumed that there are differences in response between two subpopulations. This is indicated by the different dose-response models: subpopulation 1 is less susceptible than subpopulation 2 (although it assumed that both populations have the same level of exposure). The different dose-response models lead to different individual levels of risk, with a randomly selected individual from subpopulation 1 having a higher probability of illness than a randomly selected individual from subpopulation 2. The population-level risk aggregates the results from the two subpopulations to give the number of cases in the total population. If it can be assumed that there is no difference in response between subpopulations, then stratification would not be required (Figures 5.4 and 5.6).

The examples shown here can easily be extended to recognize, for example, variability across producers, across time or area (as shown in Figure 5.3). In addition, further estimates of risk could be derived.
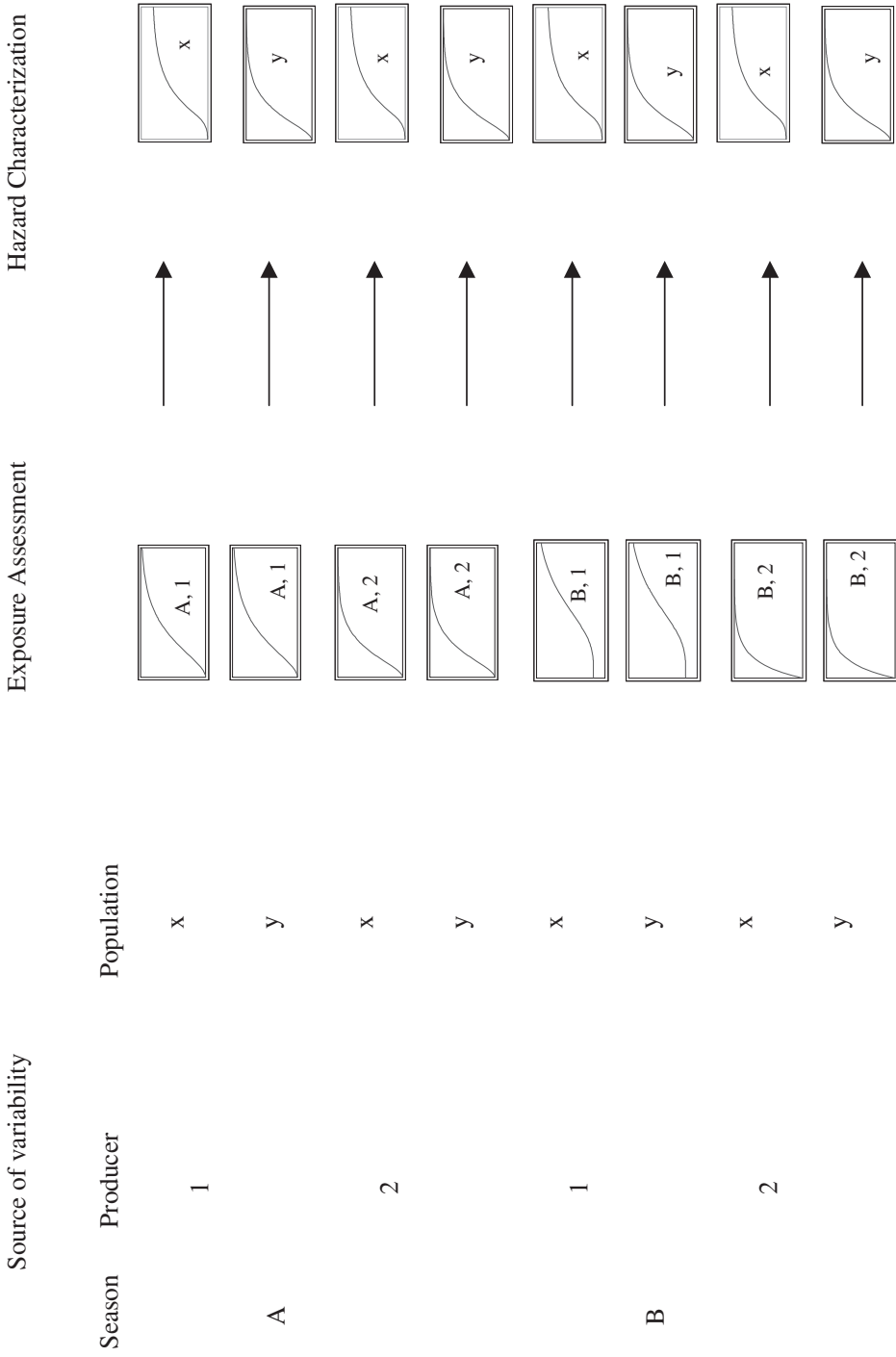
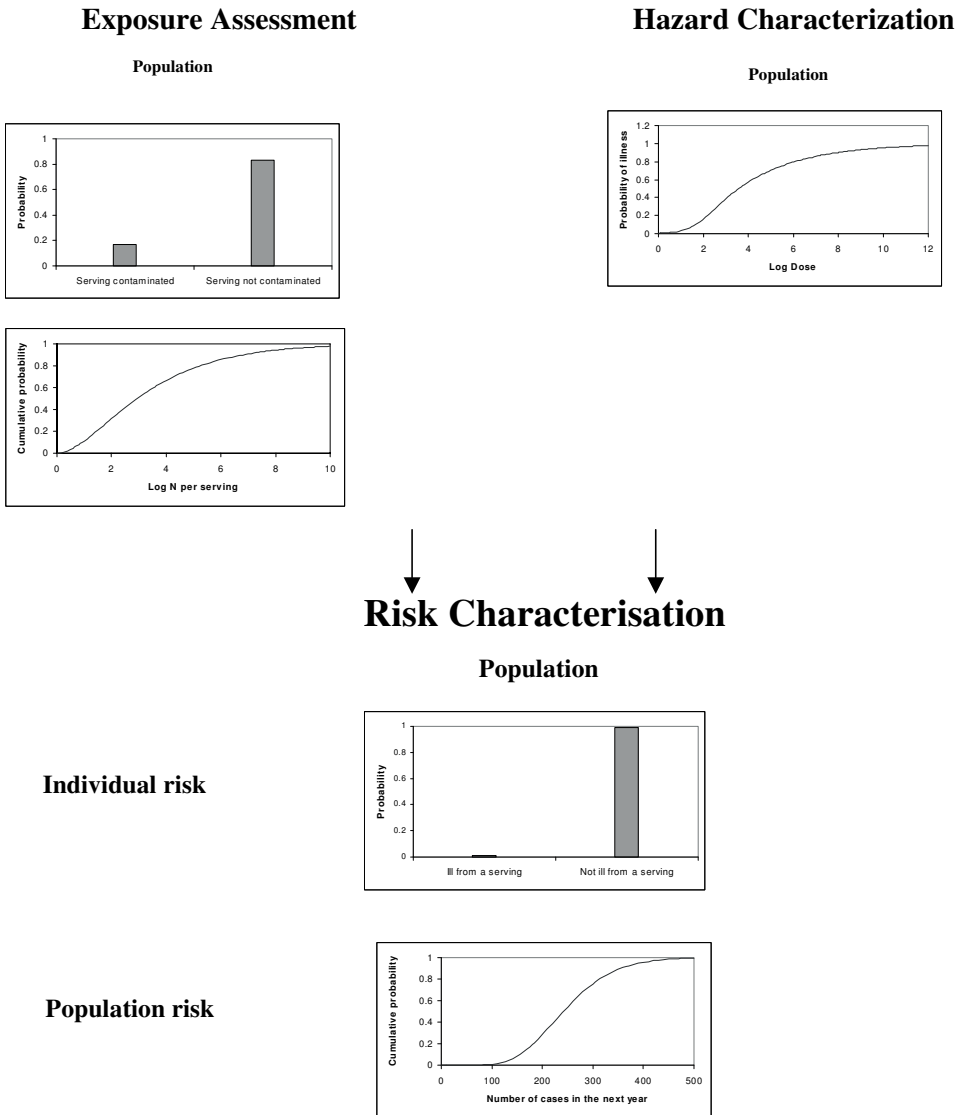**Figure 5.3** Linkage between exposure assessment and hazard characterization

Exposure Assessment

Population

Hazard Characterization

Population

Risk Characterisation

Population

Individual risk
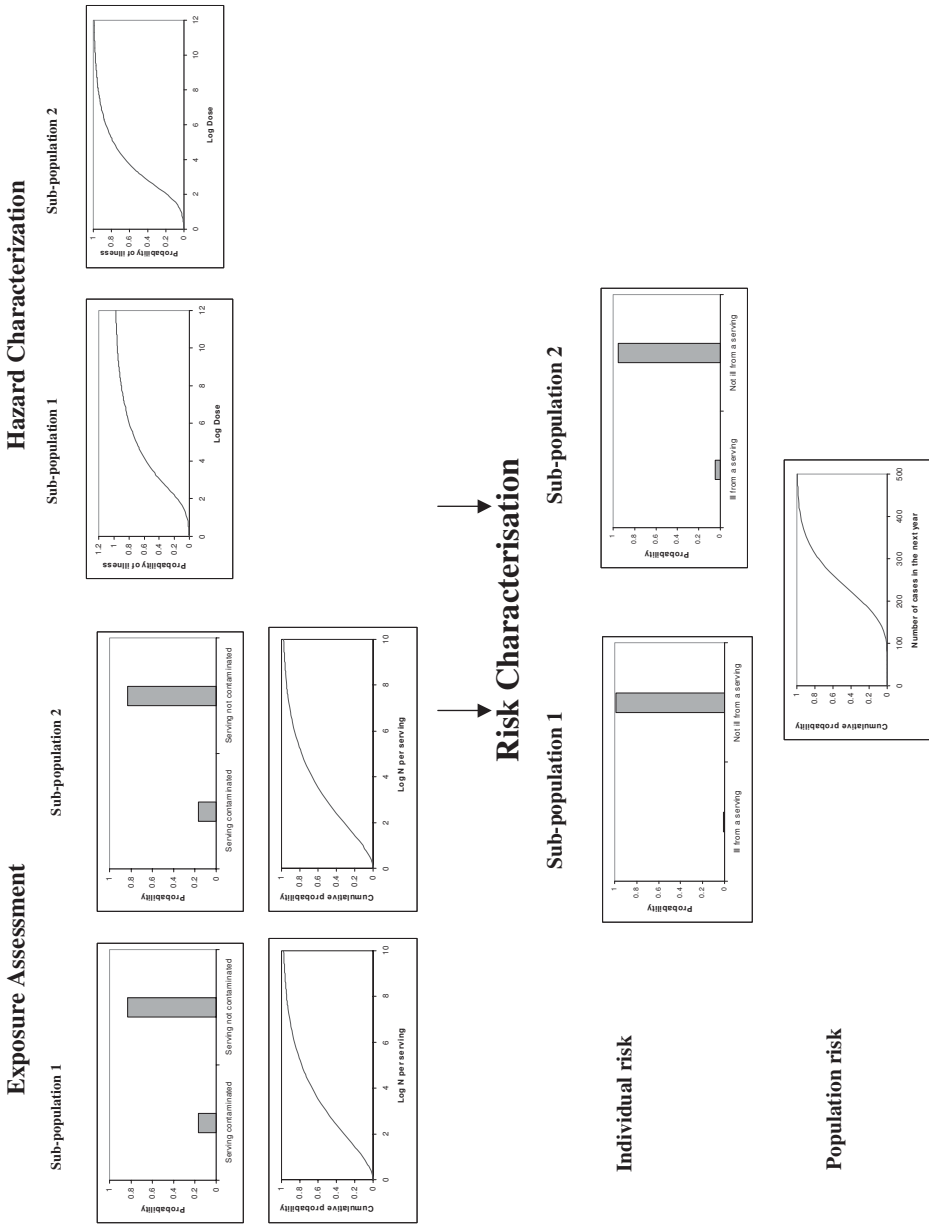
Population risk

Figure 5.4. Risk characterization without stratification of the population and uncertainty not included

Figure 5.5. Risk characterization with stratification of the population and uncertainty not included.
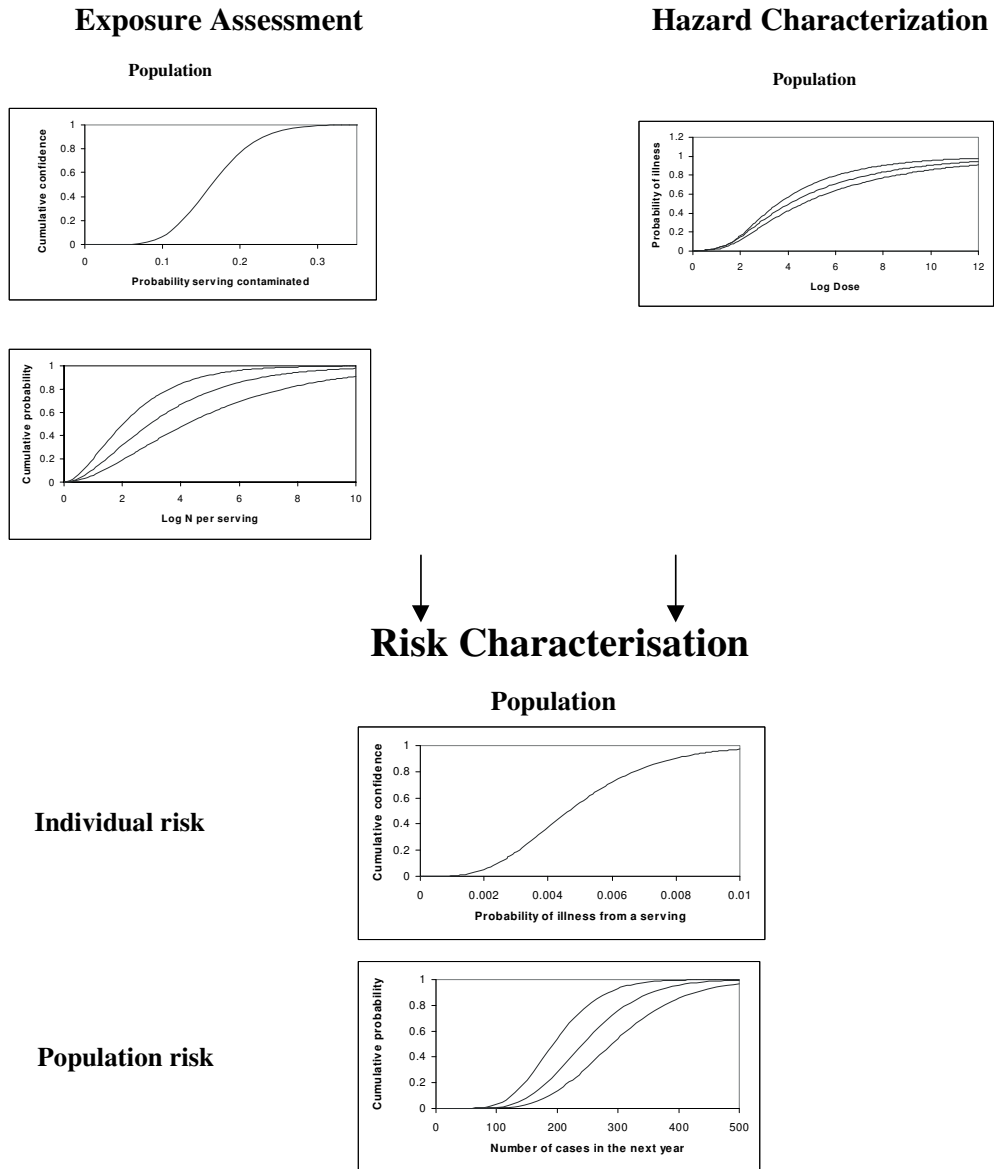
Exposure Assessment

Population

Hazard Characterization

Population

Risk Characterisation

Population

Individual risk

Population risk

Figure 5.6 Risk characterization with no stratification of the population and uncertainty included

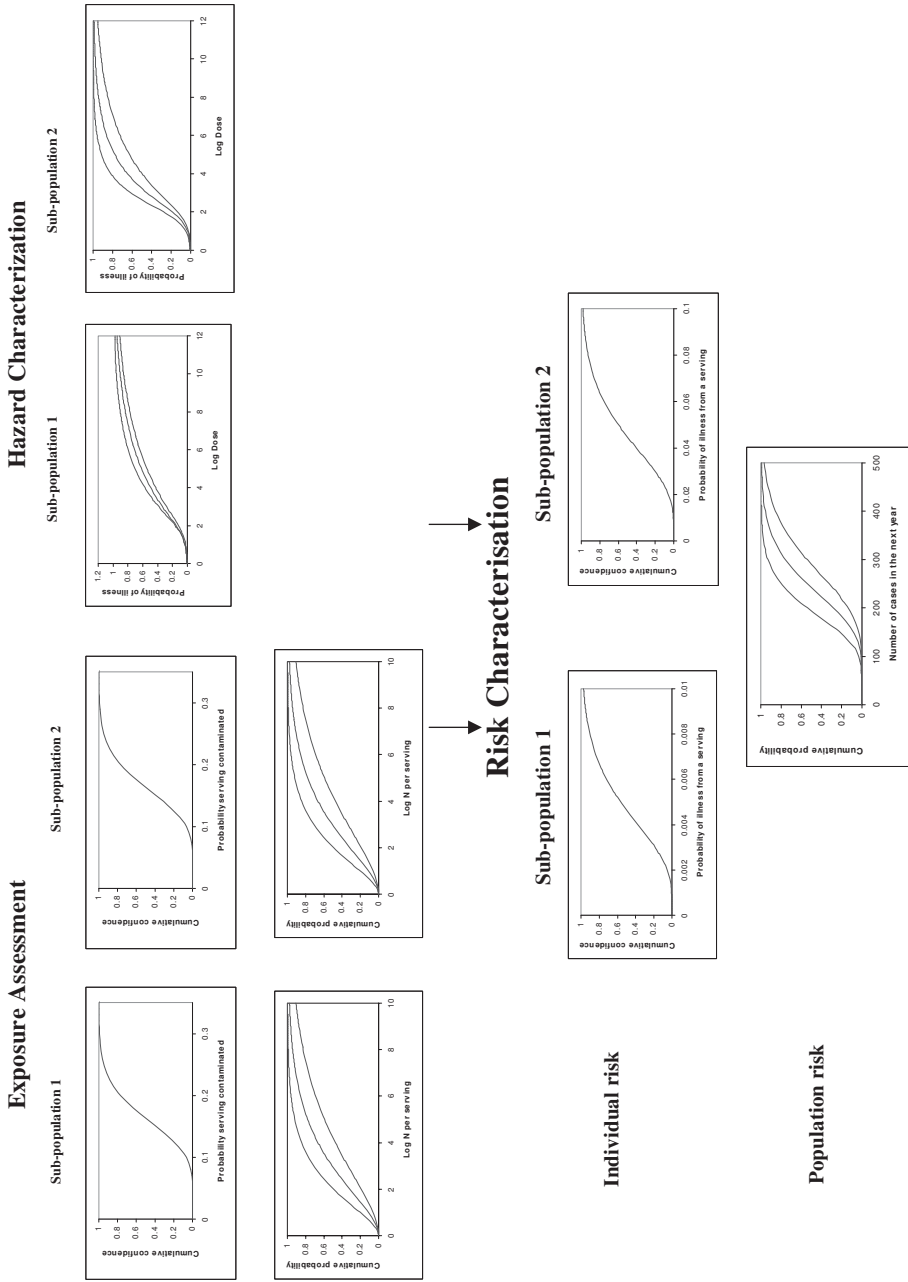**Figure 5.7** Risk characterization with stratification of the population and uncertainty included

5.6 Examples of quantitative risk analysis

5.6.1 FSIS *E. coli* comparative risk assessment for intact (non-tenderized) and non-intact (tenderized) beef (USDA FSIS, 2002)

Mechanical tenderization performed using stainless steel blades or needles translocates pathogens from the surface of intact beef cuts to beneath the surface thereby potentially shielding those pathogens from the lethal effects of heat during cooking.

FSIS wished to estimate whether blade-tenderized steak posed a significantly greater risk than its equivalent non-tenderized steak. They created a quantitative simulation model that looked at the bacterial levels on the steaks, and the change in survival of bacteria due to the extra protection that was afforded by being embedded in the meat through the tenderizing process. They then estimated the bacterial load on steaks post-cooking, and used this concentration as input into a dose-response model to estimate risk. FSIS concluded:

"The probability of *E. coli* O157:H7 surviving typical cooking practices in either tenderized or not-tenderized steaks is minuscule … 0.000026 percent (i.e. 2.6 of every 10 million servings) of steaks that are not tenderized contain one or more bacteria. For tenderized steaks, 0.000037 percent (i.e. 3.7 of every 10 million servings) contain one or more bacteria. … Differences in bacterial dose after cooking attributable to tenderized versus not-tenderized steaks are minimal at most. [The model] shows a barely discernable difference at dose levels greater than 1 between tenderized and not-tenderized steaks. The expected illnesses per serving (IPS_{EV}) for tenderized steaks is 1 illness per 14.2 million servings (7.0×10^{-8}). For not-tenderized steaks the IPS_{EV} is 1 illness per 15.9 million servings (6.3×10^{-8}). What this means is that there will be seven additional illnesses due to tenderization for every billion steak servings (7.0×10^{-8} - 6.3×10^{-8})"

This was a comparative risk assessment so the model contained only the elements that were necessary to make the comparison. Thus, the model began with the distribution of bacteria on steak prior to tenderizing, and then looked at the difference in human health risk posed by the same steak under different processing, and did not need to consider any factors involved in the rearing and slaughtering of the animal.

5.6.2 FAO/WHO *Listeria monocytogenes* in ready-to-eat foods (FAO/WHO, 2004)

FAO/WHO convened a drafting group to address three questions relating to *Listeria monocytogenes* that were posed by the Codex Committee on Food Hygiene (CAC, 2000), namely:

- Estimate the risk of serious illness from *L. monocytogenes* in food when the number of organisms ranges from absence in 25 grams to 1000 colony forming units (CFU) per gram or millilitre, or does not exceed specified levels at the point of consumption.

- Estimate the risk of serious illness for consumers in different susceptible population groups (elderly, infants, pregnant women and immuno-compromised patients) relative to the general population.

- Estimate the risk of serious illness from *L. monocytogenes* in foods that support its growth and foods that do not support its growth at specific storage and shelf-life conditions.

The risk assessment did not need to complete a full farm-to-fork model to answer these questions. The questions are also not specific to a particular country or product, which would require defining the scope of the model. The team decided to focus on the level of *Listeria*

monocytogenes at retail; model the growth and attenuation from retail to consumption; and use a fitted dose-response function to estimate the subsequent risk.

The team selected four ready-to-eat foods to be reasonably representative of the many different foods available. The quantitative analysis produced the results shown in Table 5.1.

Table 5.1 Estimated risk from *Listeria monocytogenes* as used in the FAO/WHO MRA.

Food	Cases of listeriosis per 10^9 people per year	Cases of listeriosis per 10^9 servings
Milk	910	0.5
Ice cream	1.2	0.0014
Smoked fish	46	2.1
Fermented meats	0.066	0.00025

SOURCE: Adapted from Table 1 of FAO/WHO *Listeria monocytogenes* ready-to-eat risk assessment (FAO/WHO, 2004).

The risk assessment report provides a very detailed explanation of the important limitations of the quantitative analysis, and in particular the need to rely on mostly European quantitative data on contamination, and on multiple sources for prevalence estimates. Consumption data were mainly North American, and the dose-response relationship was derived from epidemiological data from the United States of America, which may not have the same exposure levels as the European data. Its summary response to the three Codex questions recognizes the caution that should be applied in interpreting the quantitative figures, by providing qualitative responses. For example (FAO/WHO, 2004):

 "(T)he risk assessment demonstrates that the vast majority of cases of listeriosis result from the consumption of high numbers of *Listeria*, and foods where the level of the pathogen does not meet the current criteria, whatever they may be (0.04 or 100 CFU/g). The model also predicts that the consumption of low numbers of *L. monocytogenes* has a low probability of causing illness. Eliminating higher levels of *L. monocytogenes* at the time of consumption has a large impact on the number of predicted cases of illness."

5.6.3 Shiga-toxin-producing *E. coli* O157 in steak tartare patties (Nauta et al., 2001)

In a risk assessment on Shiga-toxin-producing *E. coli* O157 in steak tartare patties, Nauta et al. (2001) simulated the exposure to the population in the Netherlands using a farm-to-fork Monte Carlo model. This risk assessment provided an example of integration of exposure assessment and hazard characterization with a low-level dose and an individual-level dose-response relation. The baseline model prediction of the exposure was characterized by a prevalence of 0.29% contaminated tartare patties and a mean number of 190 CFU per contaminated patty. The distribution of CFU in contaminated patties is summarized in Table 5.2. This distribution, when combined with consumption data on the probability of consumption of a steak tartare patty per person per day, results in an exposure assessment for the population in units of CFU per person per day.

The dose-response model developed for hazard characterization was based on a well documented outbreak in a primary school in Japan (Shinagawa, 1997). Data were fitted to an exponential model separately for children and adults, resulting in point estimates for the probability of infection by a single cell of r=0.0093 for children and r = 0.0051 for adults.

Table 5.2 Distribution of exposure to STEC O157 in steak tartare patties.

CFU per exposure	probability
1	63.9%
2–10	28.8%
11–100	6.3%
101–1000	0.9%
>1000	0.11%

Due to generally low levels of CFU per exposure, the exposure distribution was combined with the dose-response model in a Monte Carlo simulation by applying the single-hit model in the form $1-(1-r)^n$, with n a random sample from the exposure distribution. Risk characterization using this approach resulted in a predicted attack rate of 0.0015% infections per person per year in the Netherlands; that is, 2335 infections per 15.6 million people per year.

Note that in this example uncertainty is not quantified; only variability is incorporated.

5.6.4 FAO/WHO risk assessment of *Vibrio vulnificus* in raw oysters (FAO/WHO, 2005)

An FAO/WHO assessment of the risk of illness due to *V. vulnificus* in raw oysters was undertaken by adapting a risk model structure previously developed in the United States of America for *V. parahaemolyticus* (FAO/WHO, 2005),. This risk assessment provides an example of integration of exposure assessment and hazard characterization, when a dose-response estimated from aggregate-level data displays appreciable bias when interpreted as valid on the level of individual exposures. A principle objective in constructing the model for *V. vulnificus* was to investigate potential effectiveness of mitigations after development of a baseline model.

A dose-response relationship for *V. vulnificus* was obtained by fitting a parametric model (Beta-Poisson) to estimated population- or aggregate-level data on arithmetic mean risk versus arithmetic mean dose over groupings of the data defined by season and year. These estimated dose-response data were based on epidemiological surveillance of cases, consumption statistics and model-based estimates of *V. vulnificus* density. The resulting dose-response model fit was interpreted as an empirical fit. By construction, the result of integrating, or recombining, the derived dose response with the baseline exposures used to develop it should be equal, on average, to the population-level mean risks on which it was also based. This, however, was not the case when the estimated dose response was interpreted as applying at the level of individuals as well as on the level of the groupings from which it was derived: an apparent consequence of the effect of cross-level bias.

The magnitude of the difference between risk predictions obtained under these two alternative interpretations of the dose response is shown in Table 5.3. Assuming that the fitted population-level risk versus dose relationship applied at the individual level resulted in predictions of risk that were consistently lower (by up to 75%) than the epidemiological estimates of mean risks. The predictions of risk obtained based on an aggregate-level interpretation of the dose response were necessarily more consistent, on average, with the epidemiological estimates of mean risks used to obtain the dose-response fit. Consequently, this interpretation was used for risk characterization.

Table 5.3 Mean risk of illness due to *Vibrio vulnificus* per serving or exposure.

Season	Estimated data based on case reports and consumption statistics	Based on fitted dose response interpreted as an individual-level risk versus dose relationship	Based on fitted dose response interpreted as mean risk versus mean dose relationship
Winter	1.4×10^{-6}	5.1×10^{-7}	1.1×10^{-6}
Spring	2.8×10^{-5}	1.7×10^{-5}	3.4×10^{-5}
Summer	4.9×10^{-5}	2.8×10^{-5}	3.9×10^{-5}
Autumn	1.9×10^{-5}	5.1×10^{-6}	2.3×10^{-5}