Quantitative microbial risk assessment: methods and quality assurance

Kwantitatieve microbiologische risico-evaluatie: methoden en kwaliteitsgarantie

1-I. Boone, 1*Y. Van der Stede, 2M. Aerts, 1K. Mintiens, 3G. Daube

1Veterinary and Agrochemical Research Centre (VAR), Coordination Centre for Veterinary Diagnostics, Groeselenberg 99, 1180 Brussels, Belgium
2Hasselt University, Interuniversity Institute of Biostatistics and Statistical Bioinformatics. 3590 Diepenbeek, Belgium
3University of Liège, Faculty of Veterinary Medicine, Food Science Department – Microbiology Section, Sart-Tilman, Bât B43bis, 4000 Liège, Belgium

*yves.vanderstede@var.fgov.be

ABSTRACT

Quantitative microbial risk assessment (QMRA) is used to estimate the risk level of pathogens along the food chain and to support management decisions for the reduction of food safety risks. The degree of credibility that can be attached to risk assessment results depends largely on the quality and quantity of the data, the model structure and the assumptions made. Quality Assurance (QA) in QMRA is defined as the structure that ensures that all the steps in the risk evaluation process are scientifically based so that the policy questions being posed can be answered. Whereas sensitivity analysis and scenario analysis are generally applied in QMRA, formal methods for the evaluation of data quality, the critical evaluation of assumptions, structured expert elicitation, the checklist approach and peer review are rarely used in QMRA, even though they would improve the transparency of the risk analysis process. An overview of QA methods for QMRA is presented. The degree of implementation of these methods should be proportionate to the stakes of the risk management questions and should be discussed in consultation between the risk assessors and the risk managers.

SAMENVATTING

Kwantitatieve microbiologische risico-evaluatie (QMRA) wordt gebruikt als een beleidsondersteunende methode met het oog op de reductie van voedselveiligheidsrisico’s. De geloofwaardigheid van de conclusies gebaseerd op een risico-evaluatie zijn onlosmakelijk verbonden met de kwaliteit en kwantiteit van de data, de modelstructuur en de gemaakte assumpties. Kwaliteitsgarantie wordt gedefinieerd als de structuur die ervoor zorgt dat alle stappen in het risico-evaluatieproces wetenschappelijk onderbouwd zijn zodat de beleidsvragen beantwoord kunnen worden. Terwijl gevoeligheids- en scenarioanalyses algemeen toegepast worden in QMRA, worden formele methoden voor de evaluatie van datakwaliteit, de kritische beoordeling van aannamen, gestructureerde expertbevraging, checklists en peer review slechts zelden gebruikt, alhoewel deze de transparantie van het risicoanalyseproces ten goede zouden komen. De mate van het toepassen van deze kwaliteitsgarantiemethoden dient afgestemd te worden op het belang van de beleidsvragen en in onderling overleg tussen analisten en beleidsverantwoordelijken besproken te worden.

WHAT IS QUANTITATIVE MICROBIAL RISK ASSESSMENT?

Definition and types

Quantitative microbial risk assessment (QMRA) is a scientifically based process used to quantitatively estimate the adverse health effects resulting from exposure to micro-organisms. In particular, QMRA is used to tackle food safety problems caused by the intake of contaminated food products (Voysey and Brown, 2000). Risk assessment, which is an integral part of the risk analysis process, is interrelated with two other components: risk management and risk communication (Figure 1).

QMRA is able to provide policy makers with a scientific basis for selecting from among the various appropriate intervention options, for determining the

---

**Figure 1. Framework of the risk analysis process adopted by the Codex Alimentarius Commission (Adapted from: FAO/WHO, 2006).**
risk-based food safety targets, and for establishing the levels of protection between countries. It thereby plays an important role in international trade (Havelaar et al., 2004; Nauta and Havelaar, 2008).

Historically, the methodology for microbial risk assessment (MRA) was derived from chemical risk assessment, but there is an essential difference between the two. QMRAs have to take into account the fact that microbial organisms – unlike chemical agents – multiply and/or can be inactivated or die within or on food products during the consecutive phases of the entire food chain and beyond. The highly variable dose-response relation between the biological agent and the adverse health effects, which is due to the specific characteristics of the micro-organisms involved and the existence of susceptible sub-populations within the target human and/or animal populations, makes the implementation of QMRAs a very challenging undertaking (Voysey and Brown, 2000).

QMRAs are carried out by developing a model which is a schematic or mathematical representation describing a food safety problem in as great a detail as possible. These models integrate information from various sources, including published and unpublished scientific studies, monitoring data, surveillance data and laboratory diagnostic data. The data can originate from disease outbreak investigations, food consumption surveys, national and international risk assessments, and so on. In addition, expert opinion is often used to fill in data gaps in risk models (FAO/WHO, 2006).

Quantitative microbial risk assessment (QMRA) is the approach that is most advanced in terms of complexity and resource requirements. As opposed to qualitative microbial risk assessments, which express outputs in descriptive terms, QMRAs require quantitative data in order to provide numerical expressions of risk, which allow for the quantification of uncertainty and variability. QMRAs can be classified as either deterministic or stochastic. In the deterministic approach, the variables are represented by single-point estimates, whereas in the stochastic approach, probability distributions are used to describe variables. The stochastic approach is generally regarded to be the most able to adequately represent or mimic the real world, even though it is often complex, data-demanding and difficult to generate (FAO/WHO, 2006).

**Guidelines for QMRA**

Very often the QMRA process is subject to constraints due to poor data quality, the limited amount of...
time and resources, the assumptions made, the deficiencies in the model structure, and the interpretation of the results. In this regard, the Codex Alimentarius (1999), which states general principles for conducting microbial risk assessment, can serve as a highly useful basis for a quality assurance framework.

A formal QMRA must be subdivided into four phases, namely: hazard identification, exposure assessment, hazard characterization and risk characterization (Figure 2). The FAO/WHO has developed guidelines for carrying out three segments of the microbial risk assessment process: the hazard characterization (FAO/WHO, 2003), the exposure assessment (FAO/WHO, 2008), and the risk characterization (FAO/WHO, 2009).

Although QMRA must have a scientific basis, value judgments, choices and assumptions are often unavoidable. Transparency should play a major role throughout the risk assessment process. Constraints, assumptions and value judgments should be documented systematically, and the risk estimates should contain a full description of the uncertainties, including their location within the risk assessment process. The data needs to be of sufficient quality, and the influence of the estimates and assumptions used in the risk assessment on the final outcome should be evaluated. In addition, there must be clear communication of the purpose and output of the risk assessment, as well as of the interaction between the risk assessors, the risk managers and the stakeholders (Codex Alimentarius Commission, 1999).

Examples of quantitative microbial risk assessments

Up to now, few QMRAs cover the entire food production chain encompassing the primary production, processing, distribution, food preparation and consumption stages, which are typically modeled as modules, with the results of one module being exploited as inputs for the following module. Such large-scale risk assessments, termed farm-to-fork risk assessments, are usually commissioned by environmental, veterinary, public health or food safety authorities, and are carried out by multidisciplinary consortia. The choice of the modeling techniques applied should be according to the problem which has to be modeled, and it depends also on the available data and expertise. Among the modeling approaches, the Modular Process Risk Modeling (MPRM) method (Nauta et al., 2001) is designed to model the transmission of micro-organisms along the food pathway by breaking down the pathway into consecutive modules and then modeling the basic microbial processes that take place in each module, such as growth, inactivation (pathogen-related) and the production processes (mixing, partitioning, removal and cross-contamination). An inventory of representative farm-to-fork QMRAs with an indication of country, agents and quality assurance methods applied is shown in Table 1. These QMRAs were principally related to three food-borne bacteria: Salmonella, Escherichia coli and Campylobacter in a variety of animals and food products, and they were were carried out by North American and European consortia between 1997 and 2010.

In a review of early QMRAs, Schlundt (2000) commented that few formal QMRAs had been carried out in accordance with the Codex Alimentarius guidelines. From the QMRAs reviewed, it was not clear whether a critical evaluation of input data had taken place, and the variability and uncertainty of the data were often not described in sufficient detail. In addition, assumptions having an impact on the final result were often not clearly presented or critically evaluated. Unfortunately, even in more recent QMRAs the same drawbacks relating to the lack of a coherent quality assurance system are still frequently encountered. The purpose of this review is to present a summary of methods that make it possible to meet the general guidelines set up by the Codex Alimentarius (1999), and thereby to contribute to the quality assurance of the QMRA process. The role of the different quality assurance methods and their usefulness in QMRA are further explained in the next section.

WHAT IS QUALITY ASSURANCE?

Definition

The assurance of quality (Quality Assurance (QA)) is the framework that is provided to ensure that all tasks included in the risk assessment are executed in a technically and scientifically correct manner, and that all model-based analyses are reproducible. The aim of the QA process is to enhance the credibility of the model results (i) by ensuring the proper interaction between risk assessors and risk managers, and by clearly defining the purpose of the risk assessment; (ii) by means of rigorous validation testing against independent data; (iii) by means of uncertainty assessment and (iv) by means of independent peer review of the various stages of the risk assessment (Refsgaard et al., 2005). The criteria that can be used to determine the validity and utility of a QMRA and that are relevant for the QA process have been summarized by Lammerding et al. (2007) (Table 2).

Quality Assurance in QMRA: Why?

As QMRA is a decision support tool used by risk managers that can have an impact on a variety of different stakeholders (policymakers, funding organizations, farmers, the meat processing industry, consumers, etc.), it is essential to know whether the results provided by the risk assessment process are sufficiently relevant, robust, credible and accurate to provide an answer to the risk problem.

In order to facilitate decision making, risk assessors need to clearly explain and communicate to decision makers the level of confidence associated with the results and to report the relevant uncertainties (where are the uncertainties; how large are they?) and assumptions.
Table 1. Overview of available farm-to-fork quantitative microbial risk assessments, with indication of the quality assurance methods applied.

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Food product/animal species</th>
<th>Country</th>
<th>CEA</th>
<th>EE</th>
<th>EPR</th>
<th>MQC</th>
<th>MCA</th>
<th>MMS</th>
<th>MVA</th>
<th>NUSAP</th>
<th>PR</th>
<th>SA</th>
<th>WI</th>
<th>Remarks</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bacillus cereus</em></td>
<td>Pasteurized milk</td>
<td>NL</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No uncertainty analysis described</td>
<td>Notermans et al. (1997)</td>
</tr>
<tr>
<td><em>Campylobacter</em></td>
<td>Broiler chicken</td>
<td>NL</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Uncertainty not quantified, assessed by scenario analysis</td>
<td>Nauta et al. (2007)</td>
</tr>
<tr>
<td>spp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Havelaar et al. (2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chicken</td>
<td></td>
<td>DK</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rosenquist et al. (2003)*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IT</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discussion of assumptions</td>
<td>Calistri and Giovannini (2008)*</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>Ground beef hamburger</td>
<td>CA</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cassin et al. (1998)</td>
</tr>
<tr>
<td>O157:H7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beef</td>
<td></td>
<td>IR</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No clear risk beefmanagement question</td>
<td>Duffy et al. (2006)*</td>
</tr>
<tr>
<td>Ground minced</td>
<td></td>
<td>US</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FSIS (2001)</td>
</tr>
<tr>
<td>Steak tartare</td>
<td></td>
<td>NL</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nauta et al. (2001)</td>
</tr>
<tr>
<td><em>Salmonella</em></td>
<td>Shell eggs, egg products</td>
<td>US</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No quantitative uncertainty assessment (only variability)</td>
<td>USDA-FSIS (1998)</td>
</tr>
<tr>
<td>Enteritidis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Salmonella</em></td>
<td>Eggs, broiler chicken</td>
<td>Not country-specific</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discussion of assumptions No complete quantitative uncertainty analysis</td>
<td>FAO/WHO (2002b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. Typhimurium</em></td>
<td>Pork, bacon, mixed meat</td>
<td>UK</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Only variability modeled, no uncertainty Clear presentation of assumptions</td>
<td>Hill et al. (2003; 2008)</td>
</tr>
<tr>
<td>DT 104</td>
<td>products</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Salmonella</em></td>
<td>Danish dry-cured pork sausages</td>
<td>DK</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Alban et al. (2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. Typhimurium</em></td>
<td>Fresh minced pork meat</td>
<td>BE</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Delhalle et al. (2009)*</td>
</tr>
<tr>
<td>DT 104</td>
<td>(pure and mixed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Salmonella</em></td>
<td>Fresh minced pork meat</td>
<td>BE</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bollaerts et al. (2009; 2010), Boone et al. (2009a; 2009b; 2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slaughter and EU</td>
<td>Breeding pigs</td>
<td>EU</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Farm-to-fork risk assessment at EU level</td>
<td>VLA-DTU-RIVM (2010), EFSA (2010)</td>
</tr>
</tbody>
</table>

* no farm-to-fork QMRA

CEA Critical Evaluation of Assumptions
EE Structured Expert Elicitation
EPR Extended Peer Review / Public Review
MQC Model Quality Checklist
MCA Monte Carlo Analysis (Tier 3)
MMS Multiple Model Simulation
NUSAP NUSAP/Pedigree for data quality assessment
MVA Model Validation
PR Peer Review
SA Sensitivity Analysis
WI What-if Scenario Analysis
where are the assumptions in the model; what impact do they have?)


METHODS FOR THE IMPLEMENTATION OF QUALITY ASSURANCE IN QMRA

Uncertainty assessment

Uncertainty is defined as the lack of knowledge concerning input data, models and assumptions (EPA, 2003). Refsgaard et al. (2007) state that both subjective and objective aspects are important in assessing the degree of uncertainty, which they define as the degree of the lack of confidence that one has concerning the validity of the information obtained. A document outlining the general guidelines (i.e. not specifically for QMRA) for characterizing and communicating uncertainty in exposure assessment was released by the World Health Organisation in 2008 (WHO, 2008). The most detailed guidelines for uncertainty assessment and uncertainty communication have been developed for environmental risk assessment by the Dutch Environmental Agency (RIVM-MNP) (Janssen et al., 2003; Petersen et al., 2003; van der Sluijs et al., 2003; 2004).

Uncertainty terminology

The use of a coherent typology of uncertainties, such as that proposed by Janssen et al. (2005), is essential for a thorough uncertainty assessment. In this typology, uncertainty is interpreted as a multidimensional concept and distinctions are made between the

(Tables and figures are not included in the text.)

Table 2. Criteria determining the quality of a QMRA, with indication of methods for addressing the validity and utility criteria (adapted from: Lammerding et al., 2007).

<table>
<thead>
<tr>
<th>Criteria relevant to validity</th>
<th>Definition</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality and treatment of data</td>
<td>Relevant and timely data, criteria for inclusion/exclusion data</td>
<td>MCA, NUSAP, PR, UA</td>
</tr>
<tr>
<td>Inference of probability</td>
<td>Appropriate choice of distributions, adequate number of iterations</td>
<td>MQC, MVA, PR</td>
</tr>
<tr>
<td>Internal consistency</td>
<td>Sound logic and inference</td>
<td>MQC, MVA, PR</td>
</tr>
<tr>
<td>Appropriateness of assumptions, expert opinion</td>
<td>Soundness of assumptions</td>
<td>CEA, EE, MQC, PR</td>
</tr>
<tr>
<td>Epidemiological and biological credibility</td>
<td>Outcomes should be within plausible limits</td>
<td>MQC, MVA, PR</td>
</tr>
<tr>
<td>Transparency</td>
<td>Systematic development of the QMRA steps, indication of data used, data gaps (use of expert opinion, assumptions). Identification and communication of the uncertainty in the models, data, assumptions, what-if scenarios.</td>
<td>CEA, EE, NUSAP, MQC, SA, UA, WI</td>
</tr>
<tr>
<td>Peer review</td>
<td>Independent review of data, models, analysis</td>
<td>PR, MQC</td>
</tr>
<tr>
<td>Stakeholders involvement</td>
<td>As appropriate for data input, QMRA should reflect its scope to the segment of stakeholders (farmers, industry, public)</td>
<td>EPR</td>
</tr>
</tbody>
</table>

Criteria relevant to utility

<table>
<thead>
<tr>
<th>Definition</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addresses the risk management question</td>
<td>Clear definition of the problem formulation, application of the results of the QMRA.</td>
</tr>
<tr>
<td>Clarity for different audiences</td>
<td>Tiered series of reports for different groups ranging from very detailed to summary reports for non-technical audiences. Progressive disclosure of uncertainties</td>
</tr>
<tr>
<td>Explicit statement of limitations</td>
<td>Description of the model’s constraints (time, money, application of results)</td>
</tr>
<tr>
<td>Identification of risk-determining steps, knowledge gaps, conflicting evidence</td>
<td>Helps decision-makers to focus on the important steps. Clear statement of uncertainties in data and assumptions. Identification of data needs</td>
</tr>
<tr>
<td>Inclusion of what-if scenarios, evaluation of potential risk reduction strategies</td>
<td>Requires defining scenarios in interaction with risk managers</td>
</tr>
<tr>
<td>Applicable to stakeholders</td>
<td>The QMRA enhances knowledge of the food production processes and can inform stakeholders</td>
</tr>
</tbody>
</table>

CEA Critical Evaluation of Assumptions
EE Structured Expert Elicitation
EPR Extended Peer Review / Public Review
MQC Model Quality Checklist
MCA Monte Carlo Analysis (Tier 3)
MMS Multiple Model Simulation
MVA Model Validation
NUSAP NUSAP/Pedigree for data quality assessment
PR Peer Review
RC Risk communication
SA Sensitivity Analysis
UA Tiered Uncertainty Analysis
WI What-if Scenario Analysis
location of the uncertainty (where does the uncertainty manifests itself in the QMRA?), its nature (epistemic or knowledge-related uncertainty vs. stochastic uncertainty), its level (on a scale ranging from statistical uncertainty to scenario uncertainty to ignorance), the qualification of the knowledge base (see further, the NUSAP/Pedigree method), and the evaluation of the value-ladenness of assumptions resulting from subjective choices (see further, Critical evaluation of assumptions).

**Tiered uncertainty analysis**

Depending on the scope and the desired level of uncertainty assessment in the QMRA process, a tiered approach (Tiers 1, 2 and 3) is recommended both by EFSA (2006) and by WHO (2008). Tier 1 analysis starts with a qualitative estimate of all the uncertainties and provides a description of the most significant uncertainties and the relative magnitude of their influence on the assessment output. Tier 2 and Tier 3 are quantitative uncertainty assessment approaches. Tier 2 consists of the deterministic analysis of uncertainties. Different alternative point estimates are filled in for uncertain inputs in the assessment and their impact on the assessment outcome is calculated. The most detailed level and resource intensive type of uncertainty analysis is obtained via a probabilistic analysis of uncertainties (Tier 3). Compared to the Tier 1 and Tier 2 approaches, Tier 3 produces probability distributions as outputs. What is essential in a Tier 3 approach is the specification of probability distributions for the model inputs. Hereafter, computations will identify how the variability and uncertainty propagate through the model, resulting in the quantification of the variability and uncertainty in the output. In addition, a sensitivity analysis can be performed to assess how the variation of the output is affected by changes in the model inputs. The most common approach for performing Tier 3 uncertainty assessment includes Monte Carlo Analysis (MCA), Bootstrapping, and Bayesian analysis (FAO/WHO, 2008).

By identifying uncertainties qualitatively, deterministically and/or probabilistically, information on data gaps can be obtained. In order to take decisions, risk managers can ask for additional data collection to reduce the uncertainties.

**Systematic review**

QMRA's generally require a diversity of data sources to build a model. It is therefore good practice to make an inventory of what is known in the literature on a specific risk problem. A systematic review approach can be utilized to obtain quality data to be used as input in a QMRA. Systematic review is a rigorous and replicable method for the identification, evaluation and synthesis of scientific evidence for the purposes of addressing a specific topic (Sargeant et al., 2005). The steps in a systematic review include (i) the development of a focused study, (ii) the identification of relevant types of research using a structured strategy, (iii) the screening of abstracts for relevance to the study question, (iv) the quality assessment of the relevant literature using pre-determined criteria, (v) the extraction of data of sufficient quality, and (vi) the synthesis of data. In meta-analysis, a statistical technique is used (e.g. meta-regression) to combine results to provide a single estimate, whereby higher weights can be attributed to studies according to their study characteristics (study population, study method, sample size, sampling plan, etc.). The absence of published literature on a specific topic can serve as a motivation to initiate additional research, to contact database owners for the exchange of (unpublished) data, and/or to set up new experiments.

**NUSAP/Pedigree approach for the evaluation of data quality**

Good quality data is data that is complete, relevant and valid. A prerequisite for the evaluation of the data quality is that the data should be sufficiently documented with respect to its references, sampling characteristics (sample size, sample methods, temporal/geographical representativeness, distribution, diagnostic test characteristics, etc.) and validation status. A systematic review approach (see previous) can be helpful in this documentation process. The NUSAP/Pedigree approach is a method that provides a basis for the structured evaluation of data quality. The purpose of the NUSAP (Numeral, Unit, Spread, Assessment and Pedigree) system is to analyze the uncertainty in scientific procedures used to support decision-making (Funtowicz and Ravetz, 1990). NUSAP uniquely integrates quantitative uncertainty information (Numeral, Unit and Spread) and qualitative uncertainty information by using expert judgment (Assessment) and a multi-criteria assessment (Pedigree) of the scientific knowledge base of a risk assessment.
The pedigree assessment is the most innovative aspect of NUSAP. It introduces a set of criteria, brought together in a pedigree matrix, that capture the essential characteristics of the data, such as the proxy representation, the empirical basis, the methodological rigor and the degree of validation (Table 3).

The proxy criterion is used to evaluate the closeness of resemblance between the input parameter available from the data source and the actual variable that would be required in the model. The empirical basis criterion is used to evaluate the degree to which direct observations were used to estimate the input parameter. A higher pedigree score for the empirical basis was attributed to input parameters obtained from the field data compared with indirect, modeled data or data obtained by expert judgment. The methodological rigor refers to the norms used in the collection and checking of the data and the degree of acceptance of these norms by the peer community in the relevant discipline. Lastly, the validation criterion is used to evaluate the degree to which it was possible to cross-check the data against independent sources.

This pedigree matrix is an instrument used by risk assessors in discussing and evaluating data. The matrix can be used to attribute scores to each criterion on a discrete numeral scale from 0 (weak) to 4 (strong). By aggregating scores over the different criteria, overall pedigree strengths are obtained. Pedigree strengths can be graphically represented within a diagnostic diagram (Figure 3) representing the overall strengths of input parameters on the x-axis and the sensitivity of the input parameters (obtained, for example, by sensitivity analysis) on the y-axis (van der Sluijs et al., 2004). The two metrics taken together – strength and sensitivity – are a measure of the quality of a parameter. The position of the input parameters within the diagnostic diagram is a helpful tool for obtaining an overview of the weak and strong links within the model and can thereby lead to model improvement.

**Expert elicitation**

Expert elicitation is the process of eliciting subjective judgments from experts. It is used to provide input for QMRA when empirical data are either lacking, or of poor quality or difficult to obtain (van der Fels-Klerx et al., 2005). Since the elicitation of expert judgment involves subjectivity, it is prone to bias from the expert providing his/her judgment, as well as from the elicitor (person collecting the expert judgment) and from the elicitation protocol used, all of which may ultimately have an impact on the validity of the decisions based on a QMRA. The aim of a structured elicitation procedure is to reduce this bias as much as possible, and this requires thorough preparation (Cooke 1991; Morgan and Henrion, 1990; Slottje et al., 2008).

A structured expert elicitation involves the selection of the experts, explanation to the experts of the problem and the elicitation procedure, a clear definition of the quantity to be assessed, a discussion of the gaps in the knowledge, specification of the experts’ belief in a distribution, and the decision whether or not to aggregate the distributions of the different experts (van der Sluijs et al., 2004). A successfully structured expert elicitation also implies solid training in elicitation techniques. In veterinary science, the most common structured expert elicitation methods include the Delphi method and Cooke’s classical model (van der Fels-Klerx et al., 2002; 2005). In particular, structured expert opinion in accordance with Cooke’s classical model was used to provide input in a QMRA for *Campylobacter* (van der Fels-Klerx et al., 2005) and

<table>
<thead>
<tr>
<th>Score</th>
<th>Proxy</th>
<th>Empirical</th>
<th>Method</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Exact measure of the desired quantity</td>
<td>Large sample, direct measurements, controlled experiments</td>
<td>Best available practice in well-established discipline</td>
<td>Compared with independent measurements of the same variable over long period</td>
</tr>
<tr>
<td>3</td>
<td>Good fit or measure</td>
<td>Historical/field data, uncontrolled experiments, small sample, direct measurements</td>
<td>Reliable method common within established discipline, best available practice in immature discipline</td>
<td>Compared with independent measurements of closely related variable over shorter period</td>
</tr>
<tr>
<td>2</td>
<td>Well correlated but not measuring the same thing</td>
<td>Modeled/derived data / indirect measurements</td>
<td>Acceptable method but limited consensus on reliability</td>
<td>Measurements not independent proxy variable, limited domain</td>
</tr>
<tr>
<td>1</td>
<td>Weak correlation but commonalities in measure</td>
<td>Educated guesses, indirect approximation by rule of thumb estimate</td>
<td>Preliminary methods with unknown reliability</td>
<td>Weak and very indirect validation</td>
</tr>
<tr>
<td>0</td>
<td>Not correlated and not clearly related</td>
<td>Crude speculation</td>
<td>No discernible rigor</td>
<td>No validation performed</td>
</tr>
</tbody>
</table>

Table 3. Pedigree matrix for the evaluation of data quality (Source: Risbey et al., 2001a).
Critical evaluation of assumptions

The quality of a QMRA depends largely on the assumptions made in constructing the model. It is therefore necessary to identify these assumptions and to screen the model for hidden or implicit assumptions. A novel method for the critical evaluation of a model’s assumptions was developed by Kloprogge et al. (2005). This method starts with the identification of the assumptions (and hidden assumptions) and the prioritization of the model’s most important assumptions (or key assumptions). Hereafter the potential value-ladenness (degree of subjectiveness) of the key assumptions is assessed. Subsequently, “weak” links in the model are identified. The next methodological steps include the further analysis of the potential value-ladenness of the key assumptions. The revision of the assessment includes an evaluation of the sensitivity of the assumptions and of the effect of different choices made with respect to the assumptions. The last methodological step deals with what should be communicated on the basis of the assumptions analysis.

To promote a structured discussion about the assumptions, Kloprogge et al. (2005) incorporated the NUSAP/Pedigree approach (see above) and proposed a pedigree matrix (Table 4) containing six pedigree criteria: (i) the influence of situational limitations, (ii) the plausibility, (iii) the choice space, (iv) the agreement among peers and among stakeholders, (v) the sensitivity to view and interests of the analyst, and (vi) the influence of context on the results.

Table 4. Pedigree matrix for reviewing the quality of assumptions (Source: Kloprogge et al., 2010).

<table>
<thead>
<tr>
<th>Score</th>
<th>Influence of situational limitations</th>
<th>Criteria</th>
<th>Influence on results</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Choice assumption hardly influenced</td>
<td>The assumption is plausible</td>
<td>Many would have made the same assumption</td>
</tr>
<tr>
<td>1</td>
<td>Choice assumption moderately influenced</td>
<td>Limited choice from alternative assumptions</td>
<td>Several would have made the same assumption</td>
</tr>
<tr>
<td>0</td>
<td>Totally different assumption had there not been limitations</td>
<td>Ample choice from among alternative assumptions</td>
<td>Few would have made the same assumption</td>
</tr>
</tbody>
</table>

Table 5. Examples of checklists useful for the evaluation of quantitative microbial risk assessments.

<table>
<thead>
<tr>
<th>Checklist name</th>
<th>Risk assessment type</th>
<th>Characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risbey</td>
<td>Environmental modeling</td>
<td>Quality assistance for internal use</td>
<td>Risbey et al. (2005)</td>
</tr>
<tr>
<td>Dutch Environmental Agency</td>
<td>Environmental modeling</td>
<td>Easy-to-use web-based application</td>
<td>Petersen et al. (2003)</td>
</tr>
<tr>
<td>Macgill</td>
<td>Waterborne risk assessment</td>
<td>Scoring</td>
<td>Maegill et al. (2001)</td>
</tr>
<tr>
<td>Paisley</td>
<td>Import risk assessment</td>
<td>Comprehensive, not too detailed</td>
<td>De Vos et al. (2009), Paisley (2007)</td>
</tr>
</tbody>
</table>

for a QMRA on Salmonella in the pork production chain (Boone et al., 2009a).
among peers, (iv) the agreement among stakeholders, (vi) the sensitivity to the analyst’s views and interests. The pedigree matrix contained an additional criterion, designated as the “influence on results” criterion.

The influence of situational limitations refers to the degree to which the choice for an assumption is influenced by the limited amount of data, time, software, hardware and human resources. The plausibility criterion designates the degree to which an assumption is in accordance with the “reality”, while the choice space indicates the degree to which alternatives were available to choose from at the moment of making the assumption. Agreement among peers addresses the degree to which the choice of peers is likely to coincide with the analyst’s choice. Agreement among stakeholders addresses the degree to which the analyst’s choice is likely to agree with the stakeholders’ views. The influence of the analyst’s views, background and interests are taken into account in the criterion “sensitivity to views and interests of the analyst”.

The “influence on results” criterion does not evaluate the value-ladenness of the assumptions, but rather provides a rough indication of the influence of an assumption on the end result of the risk assessment. The pedigree matrix is used as a tool to score the assumptions for the different pedigree criteria. As for the evaluation of the quality of data, a diagnostic diagram can be used to identify weak and strong links within a risk model (Figure 3). Individual scores for the different pedigree criteria can be represented graphically either by kite diagrams (Kloprogge et al., 2005) or by pedigree charts (Figure 4) (Wardekker et al., 2008). The critical evaluation of the assumptions can be applied after the risk assessment has been carried out. It is, however, preferable also to apply it iteratively during the development of the risk assessment so that the insights gained from the assumptions analysis can be used for the improvement of the risk assessment. The method described by Kloprogge et al. (2005) was used for the first time to evaluate the assumptions in a QMRA on Salmonella in the pork production chain (Boone et al., 2010). The evaluation of assumptions is of the utmost importance in QMRAs with high policy relevance (target settings, for example, for the entire EU). The proposed method inevitably depends on expert judgment and on the composition of the expert groups making the evaluations.

**Sensitivity analysis**

Sensitivity analysis (SA) aims to assess how the variation in the output of a model can be attributed and apportioned to the different sources of variation in the model’s input parameters (Saltelli et al., 2000). SA can be used as a quality assurance method for the purpose of obtaining better insight into the model. SA is recommended when the aim is: (i) to prioritize potential critical control points in the model, (ii) to identify key sources of uncertainty and variability, (iii) to refine, verify and/or validate the model, (iv) to prioritize additional data collection or research, and (v) to develop what-if scenarios (Frey et al., 2004).

As a preparation for SA, it is essential that the QMRA model be well structured and documented, and that a clear distinction be made between inputs and outputs (Frey et al., 2003). In modular farm-to-fork QMRAs, it can be more straightforward to perform an independent SA on the output variables for the different modules separately (e.g. primary production, transport and lairage, slaughter and processing, and preparation and consumption), than to perform a SA on the model as a whole. In this modular approach, a clear one-to-one relationship between output and inputs may be more easily identified, whereas this relationship is often very hard to observe in the end output of a SA on the model as a whole (VLA-DTU-RIVM, 2010). Secondly, SA can be particularly difficult across modules, where units of interest are variable (e.g. the random selection of individual pigs in the primary production stage, the transport of a batch of pigs to the slaughterhouse, the half-carcasses and meat-cuts at the processing stage, the meat portions, etc.).

Guidance to select and apply SA methods in food safety risk assessment is provided by Frey et al. (2004). The choice of a SA method depends on its scope, applicability and the characteristics of the model. In a review of nine SA methods (Frey et al., 2003), ANOVA and classification and regression trees (CART) were considered to deal best with the simultaneous variations in all inputs, both the qualitative and the quantitative inputs, the non-linearity and the interactions. On the other hand, sample correlation coefficients (Pearson coefficients) and linear regression were judged to be the weakest with respect to application to nonlinear QMRA models, and Spearman rank coefficients were
What-if scenario analysis

What-if scenario analysis is a conditional analysis in which specific goals for risk mitigation can be established and evaluated. In scenario analysis, different alternative scenarios (compared to the baseline risk model) can be explored, along with their associated uncertainties. The best case and worst case scenarios can be interesting for decision makers, as they show those scenarios that explore the relevant extremes of input variables as compared to the baseline model (van der Sluijs et al., 2004). While what-if scenarios provide a basis for risk management, it is also a necessary quality assurance tool, since it makes it possible to explore the possibilities and usefulness of the QMRA model.

Before doing a scenario analysis, the scope and objectives of the analysis should be clearly defined through interaction between the risk analysts, the risk managers and the stakeholders, and each scenario should be transparently documented (van der Sluijs et al., 2004). Most published QMRA studies include what-if scenarios for the purpose of exploring mitigation strategies (Table 1).

Checklist approach

Checklists offer a structured tool to help modelers during the model building and quality control process of risk models (van der Sluijs et al., 2004) and are intended for internal use by risk assessors or external use by peer reviewers for the purpose of identifying (i) pitfalls in complex models, (ii) details in the model that are critical to policy choices, and (iii) value-laden choices. A comparison of available checklists for model evaluation is represented in Table 5. A checklist for quality assistance in environmental modeling developed by Risbey et al. (2005) is also helpful for the evaluation of QMRA models. The checklist contains questions related to the description of the objectives of the model and what role it can play in policy making. Other questions focus on the internal strength and quality aspects of the model inputs and parameters, the treatment of uncertainties, assumptions and robustness of the model, and whether the model output matches the requirements of the users. Finally, there are questions that focus on how the model results are communicated to and used by the risk managers, and how

Peer review

Peer review is the independent review of data, logic, scientific interpretation, models, assumptions and analysis of all steps in the QMRA process, to ensure that it meets the standards of the scientific community (Lammerding, 2007). Comments by peer reviewers can be helpful in terms of identifying biases and ignored uncertainties, reconsidering assumptions and/or modifying and improving the design of data collection and (statistical) analysis. The main objective of the peer review process is to improve the credibility and transparency of a QMRA. In determining the appropriate type and format of the peer review, the following aspects should be considered (OMB, 2004): individual versus panel review, timing and resources, scope of the review, selection and anonymity of the reviewers, public participation, and the processing of the reviewer comments. Peer review is recommended from the early stages of the risk assessment process onwards, such as when determining which input data and model to use. The selection of peer reviewers is a challenging task,
as most QMRAs are carried out by a multidisciplinary team. Therefore, experts from different disciplines should be involved in the peer-review process, such as statisticians, veterinarians, microbiologists, epidemiologists, and medical doctors. When necessary, economists and social scientists can be involved too. To allow for peer review, QMRAs should be transparently documented, and the reviewers should have access to all the data and models. Checklists can offer a standardized format as a support tool for the review process. The three QMRAs presented in Table 1 all mentioned that external peer review had been carried out. These included two QMRAs for Salmonella on eggs and broiler chicken, and a QMRA for Salmonella in slaughter and breeding pigs, commissioned by the USDA-FSIS (1998), FAO/WHO (2002a) and the EFSA (2010), respectively. The greatest limiting factor of peer review is the time and resources one is willing to allocate, especially when quick decisions are required for high-stakes decision problems.

**Model verification**

Model verification is defined as the process of verifying that the mathematical expressions, the definitions of the data inputs, and the logic of the model are correct and correctly implemented. It involves checking the correctness of the model formulation, the inputs, and the internal consistency of the model, and it should precede model validation (see Model validation). Model verification is facilitated when the data, model structure, methods, tools and assumptions are clearly documented (FAO/WHO, 2009).

**Model validation**

Model validation consists in verifying whether a model corresponds with the reality and is fit for its purpose. Model validation includes conceptual validation (the model represents accurately the system under study), the validation of algorithms (the model concepts have been translated adequately into mathematical formulas), the software code validation (the mathematical formulas have been correctly implemented in computer language), and the functional validation (checking the model with independent observations). A model is said to be validated when there is a close match between the model output and independent validation data. In many QMRAs, validation or even partial validation is difficult to achieve due to the lack of data or comparable independent data. As an alternative to model validation when independent validation is scarce or lacking, screening procedures and sensitivity analysis can be applied to identify the most important inputs, uncertainty assessments, and multiple model comparisons (FAO/WHO, 2009).

**Multiple model comparison**

A model is always a simplification of the reality. The mismatch between the modeled system and the reality inevitable causes model structure uncertainty. As an example, in a Danish environmental risk assessment study, five alternative models were developed by five independent consultants who used the same raw data as input for their models. The five consultants all used different approaches to answer the risk management question, which resulted in substantially different model predictions (Refsgaard et al., 2006).

Large differences between alternative models may cause confusion in the results of a QMRA and delay or hinder management decisions. On the other hand, alternative models yielding similar conclusions can support and facilitate decision-making. When time and resources are limited, it is usually better to develop a single detailed QMRA model, instead of several alternative (less detailed) QMRAs. The quality of alternative models can be assessed and compared by means of previously discussed methods, such as the checklist approach, NUSAP/Pedigree, critical evaluation of assumptions, uncertainty analysis, sensitivity analysis and scenario analysis.

**Quality of documentation and risk communication**

Clear documentation of all stages of the QMRA is essential. This should include a clear representation of the strengths and limitations of the model (data quality, critical assumptions, model structure, uncertainties), and information on how the quality assurance has been dealt with. In turn, the implementation of the different quality assurance methods (e.g. peer review, NUSAP, etc.) also depends on the clarity of the documentation of the risk assessment process, the description of the data and assumptions, etc.

The way in which the results of a QMRA is documented should be adapted to different target audiences (analysts, stakeholders, decision-makers) using the progressive disclosure of information approach (PDI) (Kloprogge et al., 2007). This implies that a full technical document with all model details for risk assessors should be complemented with a less technical report that is comprehensible for decision-makers and stakeholders. Special attention should be focused on the documentation of the uncertainties and assumptions. For guidelines on the contents, style and degree of the uncertainty information at different PDI layers (Kloprogge et al., 2007). The clarity of the information can be improved by using graphics (tables, charts). For example, the quality of the data and the assumptions can be represented with kite diagrams, pedigree charts and diagnostic diagrams (Figures 3 and 4).

Jargon should be avoided for risk managers (Kloprogge et al., 2007) and emphasis should be put on the implications of the uncertainties for policy advice, while the uncertainties should be documented in detail (probability density functions, nature, extent and sources of uncertainty) only for the risk assessors.
CONCLUSIONS AND RECOMMENDATIONS

There is a need for an overall comprehensive and harmonized set of guidelines for implementing a quality assurance framework in QMRA. This need is especially great for high stakes QMRAs. For this purpose, it would be beneficial to develop guidelines for QA at the EFSA and FAO/WHO levels. For example, the guidelines of the FAO/WHO (2008; 2009) could be kept up to date and complemented with newly developed state-of-the art QA methods and information on the available software and references.

QA of QMRA should include a critical evaluation of the data, the methods, the assumptions, the output and the associated uncertainties; ideally, it should also be peer-reviewed, and the results should be validated, if possible. Up to now, some QA methods have not yet become widespread in QMRA (e.g. formal evaluation of data quality, critical evaluation of assumptions, structured expert judgment, etc.). This can be explained by the fact that these methods are novel and/or still under evaluation, and/or that there is a lack of time, resources and expertise. In addition, clear guidelines for risk communication should be developed for QMRA, in accordance with the progressive disclosure of information (PDI) principle.

A QA system in QMRA is beneficial for the risk assessors, the risk managers and the stakeholders. It is beneficial for the risk assessors because it facilitates model improvement, because identified knowledge gaps can lead to the inclusion of more realistic assumptions, and because it focuses new research where it is really needed. Risk managers can be more confident in decision making when the results of a QMRA are backed by a QA system. Further research is needed to empirically investigate the effects of a QMRA QA system on risk management decisions.

ACKNOWLEDGEMENTS

This review was financially supported by the Belgian Federal Public Service of Health, Food Chain Safety, and Environment research program (R-04/003-METZOO) ‘Development of a Methodology for Quantitative Assessment of Zoonotic Risks in Belgium Applied to the “Salmonella in Pork” Model’. The partners in the METZOO research consortium are the Veterinary and Agrochemical Research Centre (VAR), the Schools of Veterinary Medicine of both Liège and Ghent Universities, the Institute for Agricultural and Fisheries Research (ILVO), the Federal Institute for Public Health and the Center for Statistics of Hasselt University.

REFERENCES


EFSA (2010). EFSA Panel on Biological Hazards; Scientific opinion on a quantitative microbiological risk assessment
of Salmonella in slaughter and breeder pigs. The EFSA Journal 8(4), 1547.


Petersen, A.C., Janssen, P.H.M., van der Sluijs, J.P., Risbey,


