EPA-820-R-14-009

Microbiological Risk Assessment (MRA) Tools, Methods, and Approaches for Water Media

December 2014

Office of Science and Technology
Office of Water
U.S. Environmental Protection Agency
Washington, DC 20460

Notice

The tools and procedures set forth in this document are intended to describe the United States (U.S.) Environmental Agency's (EPA) approach for conducting or revising microbial risk assessments to protect human health from exposure to water-based media. They are also intended to serve as guidance to EPA and EPA contractors for conducting microbial risk assessments.

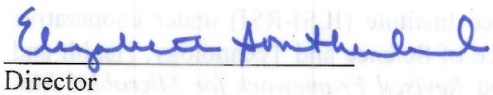
This document has been reviewed in accordance with Agency policy and is approved for publication and distribution.

Mention of commercial products, trade names, or services in this document or in the references and/or footnotes cited in this document does not convey, and should not be interpreted as conveying, official EPA approval, endorsement, or recommendation.

Foreword

This document presents a series of tools, methods, and approaches for planning and conducting microbial risk assessments in support of human health protection for water-based media. This document provides guidance for microbial risk assessments conducted or revised by EPA or EPA contractors and should not be considered regulatory.

The tools and approaches described herein focus on conducting risk assessment for water-related media (such as microorganisms in treated drinking water, source water for drinking water, recreational waters, shellfish waters, and biosolids), but are sufficiently general to help guide the development of microbial risk assessments of pathogens that might be found in food, food products, or other media.



Director
Office of Science and Technology

Acknowledgments

Project Leader
Stephen Schaub* U.S. EPA Office of Science and Technology

U.S. EPA Technical Reviewers
Alfred Dufour U.S. EPA Office of Research and Development/National Exposure Research Laboratory
Stig Regli U.S. EPA Office of Ground Water and Drinking Water
Philip Berger U.S. EPA Office of Ground Water and Drinking Water
Sharon Nappier U.S. EPA Office of Science and Technology
John Ravenscroft** U.S. EPA Office of Science and Technology

* Principal author
** Contact

Initial draft report developed by ICF International under U.S. EPA Contract 68-C-02-009; interim draft developed by Tetra Tech Clancy Environmental, Soller Environmental, and ICF International under U.S. EPA Contract EP-C-07-036; final report developed by ICF International and Soller Environmental under U.S. EPA Contract EP-C-11-005.

Contributing authors: Audrey Ichida, Jeffrey Soller, William Mendez, Jonathan Cohen, Timothy Bartrand, Mark Gibson, Jennifer Welham, Margaret McVey, Gunther Craun, Walter Jakubowski, Cynthia McOliver, and Martha Embrey.

The International Life Sciences Institute Risk Science Institute (ILSI-RSI) under cooperative agreements with the U.S. EPA Office of Water, Office of Science and Technology, Health and Ecological Criteria Division, developed a report titled *Revised Framework for Microbial Risk Assessment* to address critical areas of microbial risk assessment (ILSI, 2000). This MRA Tools document evolved out of that report. The steering committee for the development of that report included the following:

Stephen Schaub U.S. EPA Office of Science and Technology, lead
Gunther F. Craun G.F. Craun & Associates
Alfred Dufour U.S. EPA Office of Research and Development
Charles Gerba University of Arizona
Charles Haas Drexel University
Alan Roberson American Water Works Association
Mark Sobsey University of North Carolina

External Peer Review Workgroup

This document has been subject to external peer review following the Office of Management and Budget's Peer Review Guidance (2004) and EPA's Peer Review Handbook.

Potential areas for conflict of interest have been investigated via direct inquiry with the potential peer reviewers and review of their current and past affiliations. Reviewers did not have conflicts of interest.

Peer reviewers included:

Margaret Coleman	Syracuse Research Corporation (currently Coleman Scientific Consulting)
Joseph Eisenberg	University of Michigan
Charles Gerba	University of Arizona

A subsequent draft version of this document was reviewed by EPA's Science Advisory Board – Drinking Water Committee. The committee met on September 21-22, 2009. This document has been revised in response to their comments.

Table of Contents

Tables, Figures, and Text Boxes

Abbreviations and Acronyms

AGI	acute gastrointestinal illness
AIDS	acquired immune deficiency syndrome
ANOVA	analysis of variance
ARS	Agricultural Research Service (USDA)
ASM	American Society for Microbiology
AWQC	Ambient Water Quality Criteria
CAC	Codex Alimentarius Commission (Codex)
CART	classification and regression tree
CDC	U.S. Centers for Disease Control and Prevention
Codex	Codex Alimentarius Commission
CFU	colony forming units
CWA	Clean Water Act
DALY	disability-adjusted life years
EMPACT	Environmental Monitoring for Public Access and Community Tracking
EO	Executive Order
EPA	U.S. Environmental Protection Agency
FAO	Food and Agricultural Organization (United Nations)
FDA	U.S. Food and Drug Administration
FUT2	alpha (1,2) fucosyltransferase gene
GI	gastrointestinal (tract)
HACCP	Hazard Analysis and Critical Control Point
HIV	human immunodeficiency virus
HHRA	human health risk assessment
ICR	Information Collection Rule
ID_{50}	infectious dose for 50% of the exposed population
IgG	immunoglobulin G
ILSI	International Life Sciences Institute
L	liter
LD_{50}	lethal dose for 50% of the population
LT2	Long Term 2 Enhanced Surface Water Treatment Rule
MAC	*Mycobacterium avium* complex
MCMC	Markov Chain Monte Carlo method
mL	milliliter
MLE	maximum likelihood estimation
MRA	microbial risk assessment
MRM	microbial risk management
NAS	National Academy of Sciences
NRC	National Research Council
NV	Norwalk Virus
PCR	polymerase chain reaction
PFU	plaque forming unit
PMP	Pathogen Modeling Program
RA	risk assessment
QALY	quality-adjusted life years
QMRA	quantitative microbial risk assessment
RSI	Risk Science Institute (ILSI)

RT-PCR	reverse transcriptase polymerase chain reaction
SDWIS	Safe Drinking Water Information System
SMV	Snow Mountain Agent Virus
TCCR	transparency, clarity, consistency, and reasonableness
TMDL	total maximum daily load
U.S.	United States
USDA	U.S. Department of Agriculture
VBNC	viable but non-culturable
WHO	World Health Organization (United Nations)

Executive Summary

Exposure to waterborne pathogens has long been recognized as a potential source of illness in humans. Managing and minimizing this public health threat is an important aspect of the United States (U.S.) Environmental Protection Agency's (EPA) Office of Water regulatory activities and policy development. Risk assessment is a science-based tool that can be used to help managers explore the relative merits of various management alternatives, identify important gaps in knowledge, and inform regulatory actions.

This Microbial Risk Assessment (MRA) tools, methods, and approaches ("MRA Tools") was developed to assist U.S. EPA and others in conducting MRAs—including quantitative microbial risk assessments (QMRAs[1]) focused on human health risks from exposure to pathogens. The primary audience for this document is EPA staff and contractors who are responsible for conducting and managing MRAs for pathogens that occur in water and water-related media. Thus, this document is intended to summarize MRA methods and techniques for risk assessors and scientists, and provide a compilation of information that is useful for conducting rigorous, well documented, and scientifically defensible MRAs. It is not however, intended to be a comprehensive treatise, a step-by-step protocol, nor a textbook on the topic of MRA. In addition, the MRA tools document does not address deriving water quality values for microbial indicators of fecal contamination (e.g., *E. coli*, enterococci, bacteriophage), that information is addressed by EPA's recreational water quality criteria Technical Support Materials (U.S. EPA, 2014). Although the principle medium of interest is water and water-related media (e.g., recreational waters, drinking water sources, shellfish harvesting waters, biosolids), select resources for food safety risk were also consulted in the development of this document.

This MRA Tools document should be considered flexible and amenable to modification where an Office or other user has particular requirements that may not be precisely covered in the text. Moreover, the tools, methods, and approaches described herein should be considered a modular toolbox with a broad scope. It is expected that those modular aspects that are relevant to the MRA being conducted can be used as deemed appropriate by the EPA Office conducting the assessment. It is also expected that the various EPA Offices will have different needs in terms of how an MRA is documented. For example, some Offices may have a preference that specific components of the MRA be documented in a specific section of the MRA report that may be different from the manner described herein. This MRA Tools document provides for such conditions and should be considered amenable and flexible in this regard. This document does not evaluate ongoing state-of-the-art research within the field of MRA.

Microbial risk assessments can be initiated for a variety of reasons, including but not limited to the following:

- to assess the potential for human risk associated with exposure to a known pathogen;
- to determine critical points for control, such as watershed protection measures;
- to determine specific treatment processes to reduce, remove, or inactivate pathogens;
- to predict the consequences of various management options for reducing risk;
- to identify and prioritize research needs; and
- to assist in interpretation of epidemiological investigations.

[1] For the purposes of this document, the term MRA also includes QMRAs.

Individual risk assessments for specific situations can differ significantly with respect to the questions that are addressed, the information required to address those questions, and the nature of data gaps.

This MRA Tools document is comprised of a combination of concepts from numerous published risk assessment frameworks and workshop proceedings. Although many of these frameworks and proceedings were originally developed for other applications such as food safety, there are many principles that also apply to water-related risk assessments. This MRA Tools document employs an expanded and enhanced version of the EPA-International Life Sciences Institute (ILSI) Framework for MRA. For the purposes of this document, the EPA-ILSI structure has been modified in the recognition that MRA practitioners and managers often desire flexibility in the development of MRAs, and that the National Academy of Sciences (NAS), National Research Council (NRC) chemical risk framework and the EPA human health risk assessment framework are also compatible with MRA. A common theme among frameworks is the iterative nature of risk assessment. The modeling steps in risk assessment might be repeated multiple times as the scope of the assessment is refined or as risk management questions evolve. Additional data and sensitivity analyses also require repeated iterations.

Chapter 1 provides an introduction to MRA and summarizes concepts that are used throughout this document, including the purpose and scope of this document and background information. It also includes an overview of appropriate frameworks for the conduct of MRAs, such as the frameworks developed by the NRC in 1983 and 2009, and previous EPA guidance. Chapter 2 describes problem formulation and planning and scoping; it includes factors for consideration during problem formulation and a description of how problem formulation can be used to track the risk assessment progress and process. Hazard identification is also discussed in Chapter 2 as one critical component of the problem formulation process.

Exposure assessment is discussed in Chapter 3. Subtopics within the characterization of exposure include the occurrence of the infectious disease hazard, exposure assessment, and the exposure profile (a summary of the results of the exposure characterization process). Human health effects assessment, including the dose-response assessment, is the focus of Chapter 4. Subtopics within human health effects assessment include description of health effects, dose-response relationship, and the host-pathogen profile (summary of the results of the health effects assessment). Common forms of dose-response models are also summarized. In many cases EPA risk assessment documents consider the exposure assessment and human health effects assessment as two steps within an analysis phase. The third step of the analysis phase is the risk estimation, which brings together the exposure, and health effects assessments.

Chapter 5 discusses the risk characterization phase of MRA. The topics summarized include the historical context of risk characterization within EPA. Uncertainty analysis and sensitivity analysis are discussed within the context of risk characterization. Some risk assessment documentation includes the risk estimation step (bringing together the exposure and health effects assessments) within the risk characterization phase. The risk characterization phase can be interpreted as including the risk estimation step of the analysis, or only including the evaluation and context of the results presented in the analysis phase.

A total of three appendices (A-C) are also included with this MRA Tools document, to provide interested readers with additional detail on topics that are included in this document.

1. Introduction

1.1. Purpose and Scope of this MRA Tools Document

The purpose of this document is to summarize microbial risk assessment (MRA) tools, methods, and approaches for risk assessors and scientists. It also provides a compilation of information that is useful for conducting rigorous and scientifically defensible MRAs for pathogens that occur in water and water-related media. Although the principle medium of interest is water (e.g., recreational waters, drinking water sources, shellfish harvesting waters, and biosolids), MRA resources for food safety risk were also consulted in the development of this document. This MRA Tools document is designed to be flexible and amenable to modification where a user has particular requirements that might not be precisely covered in the text. It is intended to provide a modular toolbox with a broad scope and it is expected that those modular aspects that are relevant to the MRA being conducted can be used as deemed appropriate. This document is designed for use by individuals with technical expertise (e.g., microbiologists, risk assessment modelers, public health practitioners) and risk managers.

This MRA Tools document is purposely limited in its discussion and evaluation of state-of-the-art research that is ongoing within the field of MRA, and rather focuses on mature and established practices. In some places, literature is cited for readers that are interested in exploring topics in MRA's future development. Moreover, this document does not provide instructions for conducting statistical or modeling analysis. This document provides a systematic approach for framing information to be considered, and information about conducting and documenting risk assessment. It is compatible with other well-known frameworks and information to help ensure risk characterization that is helpful and relevant for the decision-makers.

This MRA Tools document focuses on MRA as it fits into the more comprehensive framework of risk analysis—an overarching term used to describe the interaction of risk assessment, risk management, and risk communication (CAC, 2004; Figure 1). Another complementary framework is the World Health Organization (WHO) Water Quality Framework (Figure 2), which provides for broad public health-based approaches for countries to assess options for meeting public health goals.

Although the NAS, NRC developed frameworks for risk assessment, those frameworks have been primarily focused on chemical risks (NRC, 1983, 2009). Microbial interactions with hosts and the environment are different from chemical interactions, and thus microbial risk assessment differs from chemical risk assessment (see Section 1.6). This MRA Tools document was developed to accommodate those differences between chemicals and microbes, while maintaining compatibility with the overall NRC frameworks.

This document focuses primarily on risk assessment and only addresses risk management and risk communication activities to the extent that they overlap with risk assessment. However, it is important to note that risk assessment is not an effective process unless risk management and risk communication activities are also comprehensively pursued. The interagency *Microbial Risk Assessment Guideline* includes more details on risk management and risk communication (U.S. EPA/USDA, 2012).

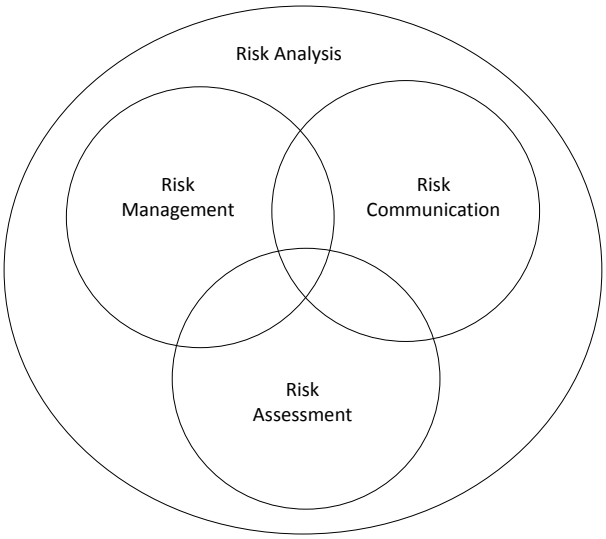


Figure 1. Risk Analysis

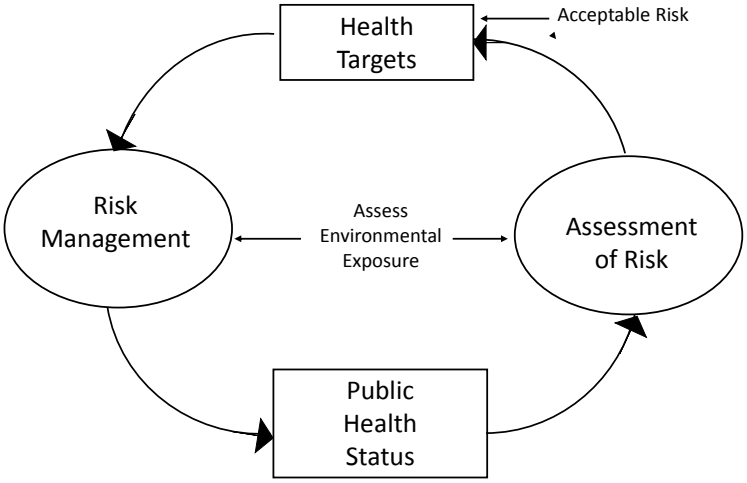


Figure 2. WHO Water Quality Framework (Source: Adapted from WHO, 2001)

One topic not addressed in this document is how EPA sets priorities for MRA or selects which MRAs to conduct. An example of one approach to priority setting is EPA's drinking water Contaminant Candidate List Classification Process.[2] The U.S. Food and Drug Administration (FDA) also has formal guidelines for determining how to set priorities for initiating MRAs (FDA, 2002).

[2] http://www.U.S. EPA.gov/ogwdw000/ndwacsum.html#ccl_cp

Microbial risk assessments can be initiated for a variety of reasons, including but not limited to the following:

- to assess the potential for human risk associated with exposure to a known pathogen;
- to determine critical points for control, such as watershed protection measures;
- to determine specific treatment processes to reduce, remove, or inactivate various pathogens;
- to predict the consequences of various management options for reducing risk;
- to identify and prioritize research needs; and
- to assist in epidemiological investigations.

Individual risk assessments for specific situations can differ significantly with respect to the questions that are addressed, the information required to address those questions, and the nature of data gaps. MRAs can be conducted to characterize the risk associated with a particular combination of a pathogen and route of exposure, "in reverse" to compute a density of a specific pathogen that would correspond to a pre-specified level of risk, or to evaluate the relative ranking of pathogen/exposure combinations. Examples of each of these approaches are referred to throughout this document.

There are some long-term goals in the MRA field that cannot yet be adequately addressed by the tools and methods that are currently available. As the field advances, this document can be expanded or modified to include new tools once they have been tested and gain general acceptance. In addition, some MRA goals have ambitious data requirements that cannot be adequately addressed at this time. Development of methods to advance MRA capabilities is a general goal of the MRA field. Some examples of possible long-term development goals for MRA are presented in Appendix A.

1.2. Development of the MRA Tools Document

This MRA Tools document is comprised of a combination of concepts from numerous published risk assessment frameworks and workshop proceedings. Although many of these frameworks and proceedings were originally developed for other applications such as food safety, there are many principles that also apply to water-related risk assessments. Some of the resources employed to develop this document are briefly summarized below and include the following:

- NAS, NRC, *Risk Assessment in the Federal Government: Managing the Process* (NRC, 1983)
- NAS, NRC, *Science and Decisions: Advancing Risk Assessment* (NRC, 2009)
- EPA and U.S. Department of Agriculture (USDA), *Microbial Risk Assessment Guideline: Pathogenic Microorganisms with Focus on Food and Water* (U.S. EPA/USDA, 2012)
- EPA Office of Water/ILSI RSI *Revised Framework for Microbial Risk Assessment* (ILSI, 2000)
- EPA *Guidelines for Ecological Risk Assessment* (U.S. EPA, 1998b)
- EPA MRA Workshops:
    - Microbiological Risk Assessment Framework Workshop: Tools, Methods, and Approaches (August 2002) (U.S. EPA, 2002c)

    - Microbiological Risk Assessment Framework: Problem Formulation Workshop (July 2003) (U.S. EPA, 2003c)
- EPA *Thesaurus of Terms Used in Microbiological Risk Assessment* (U.S. EPA, 2007a)
- EPA Office of the Science Advisor Staff Paper *Risk Assessment Principles and Practices* (U.S. EPA, 2004d)
- EPA *Framework for Human Health Risk Assessment to Inform Decision-making*, External Review Draft (U.S. EPA, 2012a)
- Codex Alimentarius Commission, *Principles and Guidelines for the Conduct of Microbiological Risk Assessment* (CAC, 1999) and *Principles and Guidelines for the Conduct of Microbial Risk Management* (CAC, 2007)
- Food and Agricultural Organization (FAO) and World Health Organization (WHO) *Microbiological Risk Assessment Series, No. 3, Hazard Characterization for Pathogens in Food and Water Guidelines* (FAO/WHO, 2003).
- WHO *Water Quality: Guidelines, Standards and Health, Assessment of Risk and Risk Management for Water-Related Infectious Disease* (WHO, 2001)

In 1983, in response to a request by the U.S. Congress, NAS, NRC (hereafter referred to as NRC framework) developed a framework that addressed primarily chemicals (NRC, 1983). It was developed by a committee of volunteer experts drawn from academia, government, and industry that was charged to conduct a study of institutional approaches to risk assessment within the federal government. The NRC committee's report underwent extensive peer review and continues to be widely cited and used in the chemical risk assessment community. The framework has also served as a template for the development of numerous subsequent risk assessments and risk assessment frameworks. In 2009, the NRC Committee on Improving Risk Analysis Approaches Used by the EPA issued a report that further developed the original 1983 framework by expanding on problem formulation and risk-based decision-making (NRC, 2009). The 2009 NRC framework has the following three phases:

- Phase I: Problem Formulation and Scoping
- Phase II: Planning and Conduct of Risk Assessment
    - Stage 1: Planning
    - Stage 2: Risk Assessment (per the original 1983 NRC framework)
    - Stage 3: Confirmation of Utility
- Phase III: Risk Management

The 2009 NRC framework also recommends formal provisions for internal and external stakeholder involvement at all stages of risk assessment. In this MRA Tools document, Chapter 2 corresponds to Phase I and Phase II (Stage 1) of the 2009 NRC framework. Chapters 3, 4, and 5 correspond to Phase II (Stage 2) of the 2009 NRC framework. Chapter 5 includes some of the concepts from Phase II (Stage 3) and Phase III of the 2009 NRC framework.

The International Life Sciences Institute's Risk Science Institute (ILSI-RSI) and the EPA Office of Water developed a conceptual framework for assessing the risks of human disease following exposure to waterborne pathogens—EPA-ILSI *Framework for Microbial Risk Assessment* (hereafter called the EPA-ILSI Framework) (ILSI, 1996, 2000). The EPA-ILSI Framework follows the general structure of the EPA *Guidelines for Ecological Risk Assessment* (U.S. EPA,

1998b). This MRA Tools document is based on and refines earlier frameworks. It follows the basic structure of EPA's *Framework for Human Health Risk Assessment to Inform Decision-making* (hereafter called EPA's HHRA Framework) (U.S. EPA, 2012a). Elements of several international frameworks for food and water microbial risk assessment have been integrated to increase the harmonization with other approaches.

The EPA-ILSI Framework describes a generic approach to identifying scientific information that should be considered in attempts to quantitatively or qualitatively assess the human health risks associated with exposure to infectious agents in water. The process to develop the EPA-ILSI Framework included three workshops held in 1995, 1996, and 1999; deliberations by a 30-member working group of scientists from academia, industry, and government; and two case study quantitative risk assessments (Soller et al., 1999; Teunis and Havelaar, 1999) to test the utility and flexibility of the framework. Notably, the participants in the 1999 workshop suggested that the framework could be further revised to include a number of additional capabilities. Two specific suggestions that are integrated into this document are the inclusion of specific information on the various types of mathematical models that have been used in MRAs and methods to address time-dependent aspects of infectious disease and immunity (dynamic modeling).

To support the continued enhancement of the EPA-ILSI Framework, EPA convened two workshops, Microbiological Risk Assessment Framework Workshop Tools, Methods, and Approaches in 2002 (hereafter referred to as the tools workshop) (U.S. EPA, 2002c); and Microbiological Risk Assessment Framework: Problem Formulation Workshop in 2003 (hereafter referred to as the problem formulation workshop) (U.S. EPA, 2003c). The tools workshop identified available analytical tools, methods, and approaches that could improve qualitative and quantitative microbiological risk assessments conducted under the existing EPA-ILSI Framework. Another important objective was to identify major issues that limit the successful application of the existing framework for conducting risk assessments.

The problem formulation workshop further developed the problem formulation stage of the EPA-ILSI Framework. Results from that workshop included elaboration of the roles of risk assessors, risk managers, risk communicators, and stakeholders during the problem formulation stage; guidance for development of conceptual models; and modification of the process diagram (flow chart) for risk assessment.

One important conclusion of both workshops was that the EPA-ILSI Framework is applicable to addressing a wide variety of public health issues related to water quality and food safety. In addition to pathogen-specific analysis, risk assessments could be used to evaluate regulatory actions, evaluate groups of pathogens (e.g., viruses), and evaluate surrogates (e.g., turbidity in drinking water). However, the EPA-ILSI Framework does not specifically discuss these types of risk assessments and did not provide examples. The discussion of problem formulation in the problem formulation workshop overlapped with EPA's Science Policy Council and Office of the Science Advisor's definition of planning and scoping (U.S. EPA, 2000b, 2002b, 2004d). Although the problem formulation workshop participants envisioned problem formulation as encompassing many of the aspects of planning and scoping, for the purpose of this document, problem formulation has been defined as part of the overall planning and scoping to be consistent with other EPA risk assessment documents such as EPA's HHRA Framework (U.S. EPA, 2012a).

A *Thesaurus of Terms Used in Microbiological Risk Assessment* (hereafter referred to as the Thesaurus) was developed in parallel to this document (U.S. EPA, 2007a). The Thesaurus compiles definitions of terms from EPA sources, other U.S. Federal agencies, international guidelines, foreign governments, and several nongovernmental organizations concerned with risk assessment. Definitions in the Thesaurus were evaluated for their potential to cause confusion, such as when the same term has differing definitions depending on its application, or when similar concepts are known by different names in different disciplines. Refer to the Thesaurus for detailed definitions of specific microbial risk concepts.

Community-based cumulative risk assessment is of growing interest to EPA. For example, EPA's Workshop on Research Needs for Community-Based Risk Assessments (October 2007) described community-based risk assessment as follows:

> Community-based risk assessment is a model that addresses the multiple chemical and non-chemical stressors faced by a community, while incorporating a community-based participatory research framework and a transparent process to instill confidence and trust among community members. It has become clear that cumulative risk assessments should include both chemical and non-chemical stressors, exposures from multiple routes, and population factors that differentially affect exposure or toxicity, and in some cases, resiliency to environmental contaminants.

Although the concepts and factors presented in this MRA Tools document could be used to consider microbial risks in the context of community-based cumulative risk assessment, at present there are no examples of cumulative MRA in the literature.

EPA's Office of the Science Advisor's Staff Paper on *Risk Assessment Principles and Practices* reviews EPA's chemical risk assessment practices across the agency (U.S. EPA, 2004d). It discusses general risk assessment topics such as conservatism, default assumptions, uncertainty, variability, and information gaps. The discussion of general topics is also applicable to MRA.

The Codex Alimentarius Commission (Codex or CAC) was created by the United Nations/FAO and WHO to develop food standards, guidelines, and related texts such as codes of practice under the Joint FAO/WHO Food Standards Programme. Codex follows an eight step Elaboration Procedure for drafting, amending, and adopting standards and guidelines. In the final step of the elaboration procedure, documents are adopted by the Commission and sent to the governments of the participating countries for acceptance. Codex adopted *Principles and Guidelines for the Conduct of Microbiological Risk Assessment* (hereafter referred to as Codex MRA Guidelines) (CAC, 1999) and a companion document, *Principles and Guidelines for the Conduct of Microbial Risk Management* (hereafter referred to as Codex Microbial Risk Management [MRM] Guidelines) (CAC, 2007). Although mainly applicable to food safety risk assessment, the Codex MRA Guidelines, and MRM Guidelines, contain many principles that also apply to water safety risk assessments because important waterborne pathogens can also contaminate foods and food products.

The FAO/WHO *Microbiological Risk Assessment Series, No. 3, Hazard Characterization for Pathogens in Food and Water Guidelines* (FAO/WHO, 2003) and the FAO/WHO *Microbiological*

Risk Assessment Series, No. 17 Risk Characterization of Microbiological Hazards in Food: Guidelines (FAO/WHO, 2009), are overall frameworks that include summaries of strengths and limitations of outbreak investigations, surveillance and annual health statistics, volunteer feeding studies, biomarkers, intervention studies, animal studies, *in vitro* studies, and expert elicitation. Elements adapted from the EPA-ILSI Framework are discussed in detail.

The WHO document *Water Quality: Guidelines, Standards and Health, Assessment of Risk and Risk Management for Water-Related Infectious Disease* (hereafter referred to as WHO Water Quality Guidelines) (WHO, 2001), is intended to harmonize[3] the process of development of guidelines and standards for drinking water, wastewater used in agriculture and aquaculture, and recreational water environments. The series of reviews in the WHO Water Quality Guidelines address the principle issues of concern linking water and health to the establishment and implementation of effective, affordable, and efficient guidelines and standards.

Concurrent to the development of this MRA tools document EPA and USDA developed and published the interagency *Microbial Risk Assessment Guideline: Pathogenic Microorganisms with Focus on Food and Water* (U.S. EPA/USDA, 2012). The interagency guideline is a useful resource, in particular for chapters on risk management and risk communication, which are beyond the scope of this document.

1.3. MRA Framework

This MRA Tools document considers the factors identified in the EPA-ILSI Framework for MRA (ILSI, 2000). The basic framework illustrated in Figure 3 is the same as EPA's HHRA Framework (U.S. EPA, 2012a) and is based on the NRC's *Science and Decisions: Advancing Risk Assessment* (NRC, 2009). MRA practitioners and managers often desire flexibility in the development of MRAs, so the other frameworks listed in Section 1.2 may also have helpful information and should be considered compatible with this MRA tools document.

As shown in Figure 3, the initial stage in conducting risk assessment focuses on carefully describing the task to be completed; it includes the planning and scoping and problem formulation components. The risk assessment phase includes developing the exposure and effects characterizations and integrating those results for presentation as part of the risk characterization. A key aspect of the HHRA Framework, "fit for purpose," is consideration of the usefulness of the assessment for its intended purpose, to ensure that the assessment produced is suitable and useful for informing the needed decisions. Attention to this concept is intended to assure, through focused planning and problem formulation and periodic reconfirmation during the process, that the informational needs of the risk managers will be met by the information being generated by the assessment. Rather than a separate step or final check in the process once the risk assessment is completed, an emphasis on the utility of the risk assessment occurs throughout the process. This begins with planning and scoping and includes evaluating the applicability of the risk assessment for informing risk management decisions; these evaluations may take place in several points of the iterative risk assessment process.

[3] In international law, harmonization refers to the process by which different states adopt the same laws (Stone, 2006). In this context it refers to the adoption of similar protocols.

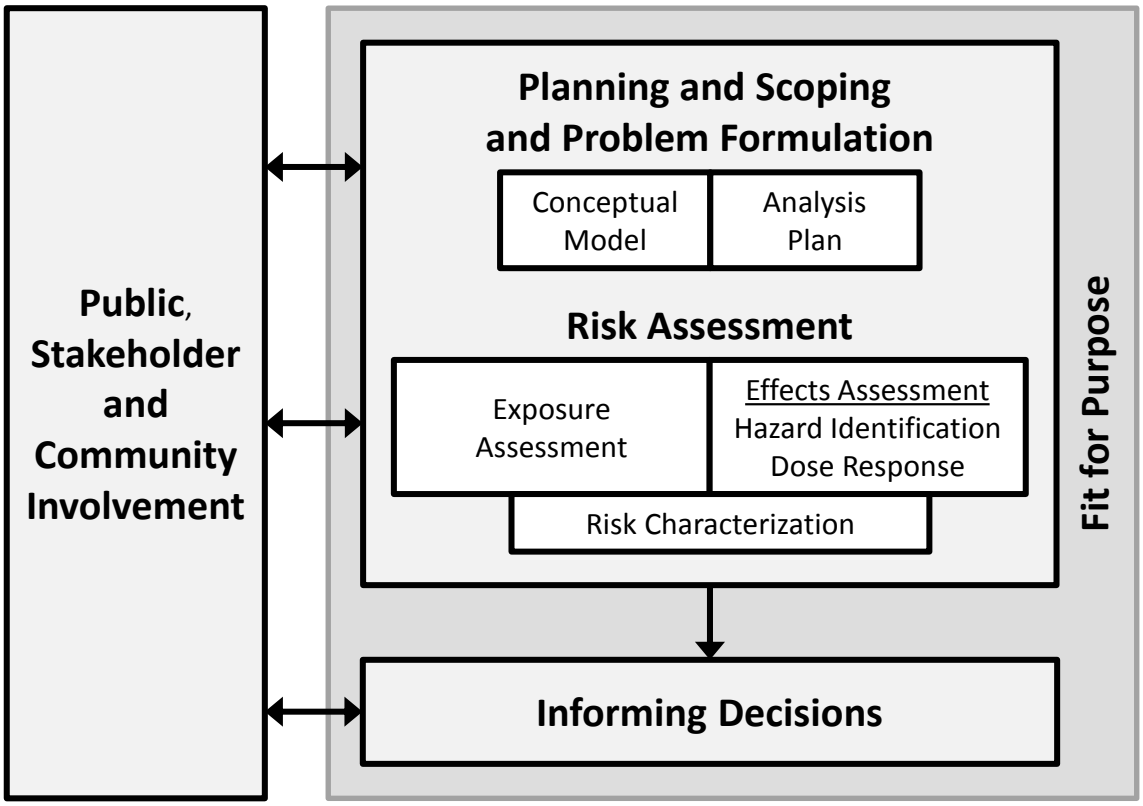


Figure 3. Framework for Human Health Risk Assessment to Inform Decision Making (Source: U.S. EPA, 2012a)

Table 1 lists factors that can be considered during the development and conduct of a microbial risk assessment. These factors are discussed in more depth throughout this document and are presented here as an overview. Not all factors will be appropriate or relevant for all MRAs and the list should not limit the addition of other factors to the risk assessment. It is helpful to provide justification when a particular factor is excluded from an MRA. Note that the factors listed in Table 1 are also referred to as "elements" or "components" in this document and can be represented by parameters in an MRA model or can be incorporated into the risk assessment in some other fashion (qualitatively). A brief summary of other risk frameworks that are consistent with this MRA Tools document is provided in Appendix B.

Table 1. Elements of Microbial Risk Assessment (Source: Adapted from ILSI, 2000; U.S. EPA, 2003c, 2003f, and 2012a)

Chapter			Elements/Factors
Planning and Scoping and Problem Formulation	Overall		Define the concern driving the RA
			Define the purpose, and objectives of the RA
			Understand the history and context within the Agency
			Define the scope of analysis
			Agree on questions the RA should answer
			Develop the conceptual model
			Develop the analysis plan (i.e., operational plan) agree on participants, roles, responsibilities, resources available, schedule, and deliverable products
			Agree on analytical approaches
	Problem Formulation	Infectious Disease Hazard Characterization	Survival and multiplication
			Resistance to control or treatment processes
			Ecology and epidemiological triad
			Virulence and pathogenicity of microorganism
			Pathologic characteristics/disease caused including host specificity
			Infection mechanisms/route of infection/portals of entry
			Potential for secondary spread
			Taxonomy/strain variation
		Host Characterization (for population of concern)	Demographics (e.g. age, size of population)
			Immune status
			Concurrent illness/medical treatment
			Genetic background
			Pregnancy
			Nutritional status
			Social/behavioral traits
Exposure Assessment Chapter	Pathogen Occurrence		Temporal distribution/frequency
			Density in environmental media
			Spatial distribution (clumping, aggregation, particles, clustering)
			Niche (ecology, non-human reservoirs, zoonotic potential)
			Survival, persistence, amplification
			Seasonality
			Meteorological and climatic events
			Presence of treatment or control processes
			Indicators and surrogates and relationships

Table 1. Elements of Microbial Risk Assessment (Source: Adapted from ILSI, 2000; U.S. EPA, 2003c, 2003f, and 2012a) (continued)

Chapter		Elements/Factors
Exposure Assessment Chapter (continued)	Exposure Assessment	Identification of media (water and shellfish for Ambient Water Quality Criteria [AWQC])
		Routes of exposure
		Units of exposure (magnitude [e.g., # of pathogen units, per volume of ingested water], duration [e.g., per day or per event], frequency [e.g., days or events compounded over a year or a life time]
		Temporal nature of exposure (whether single or multiple exposures)
		Spatial nature of exposure
		Behavior of exposed population
Effects Assessment Chapter	Health Effects	Duration of illness
		Severity of illness
		Morbidity, mortality, sequelae of illness (including acute and chronic effects)
		Extent or amount of secondary transmission
	Dose-Response	Statistical model(s) to analyze or quantify dose-response relationships
		Human and/or animal dose-response data
		Source and preparation of challenge material or inoculum in the dose-response study
		Outbreak or intervention data
		Route of exposure or administration used in dose-response study
		Equivalence of methods used (including organism type, strain, and method units) for occurrence data and dose-response study
		Characteristics of the exposed population in dose-response study (age, immune status, etc.)
		Infection or disease endpoint for the dose-response relationship (e.g., pathogen shedding, serological response, symptoms)
Risk Characterization Chapter	Risk Characterization	Evaluate health consequences of exposure scenario (risk description [event])
		Estimate the magnitude of the risk
		Conduct sensitivity analysis (evaluate most important variables and information needs)
		Summarize key issues and conclusions
		Characterize uncertainty/variability/confidence in estimates
		Address items in problem formulation
		Ensure transparency, clarity, consistency, and reasonableness (TCCR)
		Summarize assumptions including explanation of use of default values and methods
		Describe overall strengths and limitations
		Discuss how a specific risk and its context compares with similar risks

The complexity of issues surrounding the design and implementation of a microbial risk assessment requires the use of a flexible toolbox approach, in which a variety of readily available tools, methods, resources, and approaches (collectively called tools) are identified for consideration and use at different phases of the assessment. The use of a toolbox approach is

integral to using this document, although this document does not provide, nor should be inferred to provide, a comprehensive list of tools available for use in microbial risk assessment.

1.4. General MRA Concepts

This document includes various concepts and processes that broadly apply to MRAs. The following is a brief overview of some general MRA concepts.

Iterative nature of risk assessment: The risk assessment process is not linear, but flexible and dynamic (ILSI, 2000; U.S. EPA, 2012a). During any phase of the MRA process the other phases should be revisited and refined as new information and insights become available.

Transparency, clarity, consistency, and reasonableness (TCCR): Risk assessments should fulfill specific TCCR criteria (U.S. EPA, 2000b, 2012a). The TCCR criteria are summarized below.

- **Transparency**: For risk assessment to be transparent, methods and assumptions should be clearly stated and understandable to the intended audience, whether it consists of informed analysts in the field, risk managers, or the general public.
- **Clarity** refers to the manner in which the risk assessment is presented, such as writing style and the use of graphic aids.
- **Consistency** provides a context for the reader, such as whether the conclusions are in harmony with relevant Agency policy, procedural guidance, and scientific rationales, and if not, how and why the conclusions differ.
- **Reasonableness** addresses the extent to which professional judgments and assumptions are well founded, as confirmed by expert peer review. Risk characterizations should be consistent in general format, but recognize the unique characteristics of each specific situation.

Data quality: Data used in an EPA risk assessment must be consistent with EPA's *Information Quality Guidelines* (U.S. EPA, 2002a). These Guidelines build upon ongoing efforts to improve the quality of the data and analyses that support EPA's various policy and regulatory decisions and programs. They create a mechanism that enables the public to seek and obtain, as appropriate, correction of information disseminated by EPA.

Data representation: In assessing risk associated with infectious disease hazard exposures, it is usually necessary to estimate a number of parameters (quantities) in the risk models (equations) that yield numerical estimates of the probability of infection or illness. Depending on the data quality, different statistical measures (e.g., mean, median, specific percentile values) of these parameters might be appropriate.

Data variability and uncertainty: Uncertainty and variability can affect the quality and interpretation of MRA model results. Understanding, accounting for, and communicating the effects of these factors is critical in an MRA. The EPA *Exposure Factors Handbook* (U.S. EPA, 1997a, 2011) indicates that uncertainty represents a lack of knowledge about factors affecting exposure or risk, whereas variability arises from true heterogeneity across people, places, or time.

Model validation: Model validation and verification in risk assessment are general terms that are sometimes used to refer to rigorous data driven evaluation of models. However, these terms are often used interchangeably to refer to a less rigorous "reality check" that could have poorly defined validation criteria. Because validation implies different criteria in different situations, any discussion of validation should refer to how the validation was performed so that readers may properly understand the degree of rigor that the validation effort entailed. For example, one method that has been used to validate risk assessment findings is to compare the outputs to epidemiological data to determine whether the risk estimates are consistent with that which has been observed.

Risk assessment team: Risk assessment teams are multidisciplinary and may include individuals with expertise in diverse disciplines, including economics; law; engineering; the sciences (such as microbiology, epidemiology, toxicology, chemistry, and medicine); statistics; mathematics; software programming; website design; and technical writing. Although individuals may have overlapping roles, it is important that conflicts of interest between risk assessors and risk managers be avoided to maintain the scientific integrity of the process and stakeholder confidence. Risk assessment and risk management roles for risk assessment team members should be clearly defined. Note that in Figure 1 that the activities of risk assessment, risk management, and risk communication overlap. The same person may have multiple roles. In the Federal government, each agency or office will have unique considerations regarding the composition and organization of the team. The principles outlined in this MRA Tools document should be broadly applicable for a wide variety of organizational compositions.

Stakeholders: The term "stakeholders" refers to people and organizations that can shape the process or will be (or perceive themselves to be) affected by the risk assessment. Stakeholders should be involved in the Planning and Scoping in a meaningful way. At a minimum, they should be informed about the risk assessment problem, how it is to be addressed, and have an opportunity to provide comments. When stakeholders are directly affected by the proposed assessment, stakeholder comments should be sought to help team members better understand and define the problem. Stakeholders should also be informed periodically of any changes in the problem formulation.

Peer review: The role of peer review is to enhance the quality and credibility of EPA decisions by ensuring that the scientific and technical work products underlying these decisions receive appropriate levels of peer review by independent scientific and technical experts. EPA's *Peer Review Handbook* provides guidance on conduct of peer review (U.S. EPA, 2000a, 2006e, 2012b).

1.5. Microbial Risk Assessment for Decision-Making

Risk assessment is used by governments worldwide for supporting decision-making. At EPA, ecological and human health risk assessments are used to support many types of management actions, including the regulation of hazardous waste sites, industrial chemicals, and pesticides, or the management of watersheds or other ecosystems affected by multiple nonchemical and chemical stressors (U.S. EPA, 1998b, 2012a; NRC, 2009). This MRA Tools document is harmonized with EPA's *Guidelines for Ecological Risk Assessment* and EPA's *Framework for Human Health Risk Assessment to Inform Decision-making* (U.S. EPA, 1998b; 2012a).

MRA in the Federal government is always conducted in the context of supporting decision-making. Each step in risk assessment is planned and conducted within the context of the risk management issue. Risk assessment is an iterative process so that risk managers and risk assessors can work together to craft a risk assessment that answers the questions that are important to managers. In many cases the results of the first iterations of the risk assessment will inspire additional risk management questions that can be incorporated into later iterations of the risk assessment. Communication between risk assessors and risk managers is crucial for designing a risk assessment that fulfills the risk manager's needs and provides clarity for the risk managers regarding the uncertainties and caveats associated with the risk assessment.

The wide use and important advantages of risk assessments do not mean they are the sole determinants of management decisions; risk managers consider many factors. For example, decisions may be informed by a range of factors, evidence, and policy choices, such as the following (U.S. EPA, 2012a):

- Laws and Regulatory Requirements—legal mandates, flexibility and constraints.
- Economic Factors—costs, benefits and impacts of potential actions.
- Sustainability—life cycle, multimedia and long-term impacts.
- Technological Factors—feasibility, impact and range of risk management options.
- Political Factors—interactions with different branches and levels of government and the citizens that they represent.
- Public and Social Factors—susceptible population groups, nonchemical stressors and cumulative risk assessment considerations.

Reducing risk to the lowest level may be too expensive or not technically feasible. Thus, although risk assessments provide critical information to risk managers, they are only part of the environmental decision-making process (U.S. EPA, 1998b).

1.6. Factors Unique to Microbial Risk Assessment as Compared to Chemical Risk Assessment

Chemical risk assessment methods were examined for their applicability to microbial risk assessment by an EPA Office of Water workgroup. Many of the concepts developed for chemical risk assessments have parallels in MRA, but additional features have been developed to account for the differences between chemicals and microbes. Microbial risk assessments from the early 1990s identified several areas where chemicals and microorganisms differ, as noted in the sections that follow.

1.6.1. Microbial Growth and Death

Pathogens increase and decreases in number in the environment and in a host, and are variably affected by environmental and treatment factors. Different species, and even different strains within a pathogenic species, grow and die in unique patterns. In contrast, although chemicals can bioaccumulate and bioconcentrate, they are not known to multiply in the environment or in hosts.

Both chemicals and pathogens can decrease due to environmental factors, chemicals can be transformed or degrade, and pathogens can die. Not all methods used to detect and quantify microbes can distinguish between living and dead organisms; therefore, the assay method might affect data analysis when combining or comparing studies. A further complication is that several species of bacteria, including frank pathogens (e.g., *Vibrio* spp.), have been found to exist in a state called "viable but non-culturable" (VBNC). This means that, although unable to multiply on agar-medium culture plates or grow in liquid media, such cells remain functional and metabolically active (NRC, 2004). Whether pathogens in the VBNC state are infectious has not been conclusively determined (Bogosian and Bourneuf, 2001). In contrast, chemical quantification methods are generally more reproducible and able to reflect the "active" concentration of toxic agents. Microbial toxins can remain after the organism dies, and some enterotoxins are heat stable and resistant to degradation. These toxins can cause many of the symptoms of gastrointestinal (GI) tract illness.

1.6.2. Detection Methodologies

Generally, methods for detecting chemical pollutants are sufficiently sensitive to detect and quantify concentrations well below the levels that are known to have human health effects. This is not necessarily the case for pathogens. Theoretically, a single pathogenic organism can cause infection (and lead to illness). Analytical methods for detecting low levels of pathogens (e.g., one organism in 2 liters [L] of water) are not available in all cases. Although fecal indicator bacteria are useful for detecting fecal contamination, indicator bacteria do not necessarily correlate with the presence of human pathogens or public health risk (NRC, 2004). Therefore this document focuses on estimating risk from exposure to pathogens. Microbes are subject to environmental matrix effects that can cause uneven distribution that can result in consecutive measurements that differ significantly. Matrix effects can also affect the precision and accuracy of the analytical methods used to detect and quantify microbes in water.

As noted above, microorganisms in a VBNC state are also a concern for interpretation of enumeration methods. The analytical methods are probably the biggest challenge and represent the largest source of difference between chemical and microbial risk assessments. The microbial methods include microscopic techniques that do not rely on the viability of the microbe (e.g., protozoan pathogens *Cryptosporidium* and *Giardia*). The approach to establishing minimum limits of detection and practical quantification limits for microbial methods is unlike the approach taken for analytical methods used to enumerate chemical concentrations. While the assumption of one organism per volume of water sampled as a method detection limit may work for some microbial assays, it is not valid for all assays (AWWA, 2006). For example, the highly variable recovery rates for *Giardia* and *Cryptosporidium* cysts and oocysts, respectively, may be affected by the amount of processing the sample goes through before the enumeration step. In turn, this affects the reliability and reproducibility with which one oocyst or cyst can be enumerated in a sample volume. Poor reproducibility contributes to increased uncertainty as the density approaches the minimum detection limit. For these reasons, enumeration methods for microbes introduce a sufficiently high level of uncertainty that the details of those methods need to be discussed in the context of their effect on the risk assessment.

1.6.3. Genetic Diversity of Pathogens

Microorganisms are genetically diverse and allelic ratios in a population can change significantly within a few generations. In addition, microbial genomes can evolve quickly (within days or weeks) through mutation or horizontal gene transfer. Strains of the same species (e.g., *Cryptosporidium parvum*) can have multiple genotypes, potentially with different virulences for human hosts (Morgan et al., 1999; Xiao et al., 2000). Some pathogens (e.g., *Helicobacter pylori*, many viruses) behave like quasi-species, which are fluctuating populations of genetically distinct variants that co-exist within a single host (Boerlijst et al., 1996; Covacci and Rappuoli, 1998). Microbes can, thus represent a "moving target" because the distribution of strains and virulence factors can fluctuate rapidly in a given water body (Loewe et al., 2003; NRC, 2004). Variation found in the environment can also depend on different sources and types of microbial pollution. In addition microbes can acquire antibiotic resistance, which affects the range of clinical treatments that are possible and can render normally treatable illnesses life threatening.

1.6.4. Host Immunity and Susceptibility

Human hosts can have different susceptibilities to infection by particular pathogens, and levels of immunity against different pathogen species and strains may differ widely (i.e., variability among humans and variability among pathogens). Although body weight, age, and metabolic capacity differences are considered in the development of chemical criteria, genetic and acquired differences in susceptibility are not usually considered. Infection and illness due to pathogens are, in some cases, highly dependent on the immune status of the individual, which can fluctuate based on the time since the last exposure, presence of concurrent infections (e.g., human immunodeficiency virus [HIV]), and a number of other factors (e.g., life stages, gender, genetics) (Balbus et al., 2000; Parkin et al., 2003; Parkin and Balbus, 2000). For some pathogens, previous exposure may provide additional protection from that pathogen as a result of increased host immunity (Soller and Eisenberg, 2008).

1.6.5. Dose-Response Range can be Broad

The levels of pathogens required to cause infection and/or disease can vary substantially across pathogen species. Even within a particular species, those levels can vary by orders of magnitude, depending on the strain. The possible host responses may encompass asymptomatic infection, symptomatic infection (illness or disease, including chronic sequelae), and even death. Quantitative data on the exposed population's immunity and susceptibility to a pathogen and data on pathogen strain infectivity in human subgroups with differing immunity would allow the development of dose-response curves that represent a range of possible dose-response relationships. However, these types of data are not readily available. For example, although human dose-response data for six isolates of *Cryptosporidium* are available (e.g., Okhuysen et al., 2002), the data only include responses from healthy adult volunteers (for ethical reasons).

1.6.6. Secondary Transmission

Microbial infections can be transmitted from an individual to other susceptible individuals, and even to some animals. With the exception of the mother-fetus relationship, chemicals in tissues of

exposed individuals are not known to transmit to other individuals.[4] For example, in one investigation that studied person-to-person transmission of infection, the effect of rotavirus transmission within households and on the risk of infection from outside of the household was investigated through analyses of serum pairs (Koopman et al., 1989).[5] The researchers found that 17 to 20% of rotavirus infections were acquired in the household and the remainder acquired in the community. Some microbes can remain viable for days, weeks, or months, in the environment, which increases the potential for transmission. For some pathogens, humans can become asymptomatic chronic carriers and thus can infect others and contaminate food and water sources without displaying symptoms themselves for prolonged periods.

1.6.7. Heterogeneous Spatial and Temporal Distribution

Pathogens are typically heterogeneous in environmental matrices. Whereas most soluble chemicals diffuse evenly in water matrices, pathogens may clump or may be embedded in or attached to organic and inorganic particulate debris, making density determinations difficult. Although density in pipe scale and biofilms is also a problem for chemical contaminants, some pathogens can grow and/or be protected in these environments (NRC, 2004). Also, many types of pathogens occur only episodically in drinking and source waters (and in ambient waters as well) and typically can be found only during short-lived disease outbreaks (i.e., epidemics) in a community. Seasonal increases in the environment cause water or wastewater to be contaminated episodically, through breakdowns in wastewater management or water contamination controls. Therefore, contamination sources may be different for each contamination event. Seasonal fluctuations are thought to occur due to fluctuations in factors such as precipitation, temperature, nutrient availability, human activity, and livestock events (e.g., birthing season). The episodic nature of contamination makes calculation of relative sources of microbial contamination less useful than relative source contribution for chemicals.

1.6.8. Zoonotic Potential

Many, but not all pathogens also infect and amplify in animals. There is evidence that these zoonotic pathogens may change in infectivity, virulence, and the severity of disease caused in humans depending on their previous host environment. There is also evidence that some of these host-factor changes can influence subsequent infection cycles in exposed hosts (U.S. EPA, 2009a; WHO, 2004). There are six key waterborne zoonotic pathogens in the United States, *Salmonella*, *Campylobacter*, pathogenic *E. coli*, *Leptospira*, *Cryptosporidium*, and *Giardia* (U.S. EPA, 2009a).

[4] Chemicals that are on exposed individuals' clothing or skin can be transferred to household and other contacts.
[5] Serum antibodies, which are specific to different pathogen strains, indicate an immune response in an individual and are interpreted as an indicator of exposure to the specific pathogen strain for which antibodies are present.

2. Planning and Scoping and Problem Formulation

This chapter addresses the planning and scoping and problem formulation aspects of MRA. During planning and scoping the purpose of the risk assessment is defined through a dialogue between risk assessors, risk managers, risk communicators, and stakeholders. To be consistent with EPA's Science Policy Council and Office of the Science Advisor's documents on human health risk assessment, planning and scoping is considered as the broad set of activities necessary for successfully initiating a risk assessment. The overall planning and scoping considers the risk assessment within the context of overall agency resources (U.S. EPA, 2000b, 2002b, 2004d, 2012a).

Problem formulation falls within planning and scoping and can continue iteratively throughout the risk assessment process (U.S. EPA, 2000b, 2002b, 2004d, 2012a). The purpose of the problem formulation process[6] is to develop the scope of the risk assessment, taking into account management needs, Agency risk assessment policies, risk assessment tool availability, data constraints, and the nature of the decisions to be supported. At any phase in the risk assessment process, the problem formulation may be revisited.

For human health risk assessment, EPA considers the overall planning and scoping steps to be as follows (adapted from U.S. EPA, 2003f, 2003c; also see Table 1):

- defining the concern driving the risk assessment;
- defining the purpose, and objectives of the risk assessment;
- understanding the history and context within the Agency;
- defining the scope of analysis;
- agreeing on questions the risk assessment should answer;
- developing the conceptual model;
- developing the analysis plan (i.e., operational plan) agreeing on participants, roles, responsibilities, resources available, schedule, and deliverable products; and
- agreeing on analytical approaches.

It is not necessary to rigidly delineate various activities as part of planning and scoping versus problem formulation. It is sufficient to understand that problem formulation includes discussion of scientific and science policy choices related to the conduct of risk assessment while planning and scoping includes problem formulation and the operational, logistical, and budgetary planning necessary to successfully conduct the risk assessment.

2.1. Introduction to Planning and Scoping and Problem Formulation

Tasks for problem formulation include describing specific risk management questions, determining data and resource needs, performing preliminary exposure and health effects assessments, developing a conceptual model, and defining key assumptions. Forming an

[6] Note, Codex refers to this stage as "risk profile."

operational plan for conducting the risk assessment should also be accomplished during planning and scoping. If it is determined that a full risk assessment is not needed or is infeasible, information gleaned from the problem formulation stage can be used as a qualitative risk assessment or even a semi-quantitative risk assessment, and the process can, in fact, stop after the problem formulation stage. This stepwise approach can be a means of prioritizing resources and defining the scope of the overall risk assessment and to determine whether sufficient information is available to conduct a comprehensive quantitative risk assessment, if in fact, the risk management questions require a comprehensive assessment. Wooldridge and Schaffner (2008) provide guidance on qualitative risk assessment.

Identification of the nature of required inputs and outputs is necessary during problem formulation. Two general risk assessment approaches are consistent with this MRA Tools document. In the first approach, pathogen occurrence, exposure assessment, and dose-response assessment are combined to arrive at an estimated risk level. This first approach would be used, for example, to characterize the risk associated with a specific pathogen through specific route of exposure. In the second approach, which can be useful for regulatory purposes, dose-response assessment, exposure assessment, and a target risk level or risk range[7] are combined to determine a pathogen density that would provide a pre-specified level of public health protection. In the first approach, the estimated risk (e.g., event, daily, or annual risk of infection or illness) is the output; in the second approach, the pathogen density for a given exposure scenario is the output. There are also other types of risk assessments that may be consistent with this document, including the following:

- Product/Pathogen Pathway Analysis—used mainly for microbial risks in a specific food; the risk assessment models the temporal/spatial pathway a product follows through production to consumption;
- Risk Ranking—ranks risks of the same pathogen from multiple sources, or ranks risks of multiple pathogens from one source; for example see FDA-U.S. Department of Agriculture (USDA) *Listeria* risk assessment (FDA/USDA, 2003);
- Risk/Risk Analysis—compares risks between different scenarios, usually management options); and
- Geographical Introduction Analysis—used to estimate risk of introduction of disease agents through food animals or animal products (e.g., intentionally as in bioterrorism or unintentionally) to a region; for example, the risk of bovine spongiform encephalopathy ("mad-cow disease") occurring is U.S. herds due to importation of livestock from other countries.

All of these types of risk assessments may have different types of outputs and require different inputs. The information presented in this MRA Tools document should be evaluated within the context of the scope of a given risk assessment.

The WHO Water Quality Guidelines (WHO, 2001) include methods for risk assessments that have health targets, water quality targets, or a performance target that includes engineering technology (including technological approaches for small communities). In this context MRA can be used to

[7] "Target risk range" is similar to "appropriate level of protection", which is used in the World Trade Organization "Agreement on the Application of Sanitary and Phytosanitary Measures" and the Codex MRM Guidelines (CAC, 2007).

(1) provide estimates of the burden of disease, (2) establish norms and standards such as water quality, (3) assess the safety of a system against performance standards, and (4) assess health impacts. The WHO Water Quality Guidelines methodology is similar to the Codex MRM approach (CAC, 2007) in that it relates public health goals to "food safety objectives," "performance objectives," "performance criteria," and "microbiological criteria".

The WHO and many countries have adopted Water Safety Plans, which use the concept of Hazard Analysis and Critical Control Point (HACCP) (Bartram et al., 2009). Seven basic principles are employed in the development of HACCP plans that meet a stated goal, and include hazard analysis, critical control point identification, establishing critical limits, monitoring procedures, corrective actions, verification procedures, and record-keeping and documentation (NACMCF, 1997).

During the problem formulation stage, the above concepts can be discussed in the text of one or more of the suggested problem formulation components. These components, which are discussed below, include the statement of concern, statement of purpose, questions the risk assessment should address, and conceptual model narrative. For example, a risk assessment that estimates the burden of disease can compare water treatment processes, which is a technology-based performance perspective.

The problem formulation process diagram is shown in Figure 4. Note that this diagram does not include specifics about what questions can be asked or how the conceptual model should be built. However, it does show the types of information that should be collected to determine the feasibility of conducting a MRA[8].

The diagram is roughly chronological. Initially, a concern or set of concerns is identified. Those concerns can come to the attention of the Agency through various routes. The statement of concern, statement of purpose, and questions to be considered evolve throughout the problem formulation stage. These can be integrated into a risk assessment "charge." Risk managers are responsible for ensuring that appropriate problem formulation documentation is developed so that it is sufficient for the particular problem at hand. With initial information regarding the scope and questions for the risk assessment, risk assessors determine the feasibility of carrying out those plans given the available data, risk assessment tools, and time and resources. A concise conceptual model, narrative, and analysis plan are developed. A screening-level risk assessment may first be performed to determine if the risk assessment questions can be addressed without an extensive formal quantitative risk assessment. In some cases, a screening level risk assessment may be adequate for decision-making. If a formal quantitative risk assessment is desired and feasible, a more detailed conceptual model/narrative and analysis plan are developed. The problem formulation documentation can be used to assist risk managers with policy decisions that are needed to define the scope of the risk assessment. For example, risk assessors can outline options for risk managers to consider. Because risk assessment is iterative by nature, aspects considered during the problem formulation may need to be revisited multiple times as new information and/or data become available. During problem formulation, the risk assessment options that are considered, the options that are chosen, and the justification for those decisions, should be carefully tracked and documented.

[8] For the purposes of this document, the term MRA also includes QMRAs.

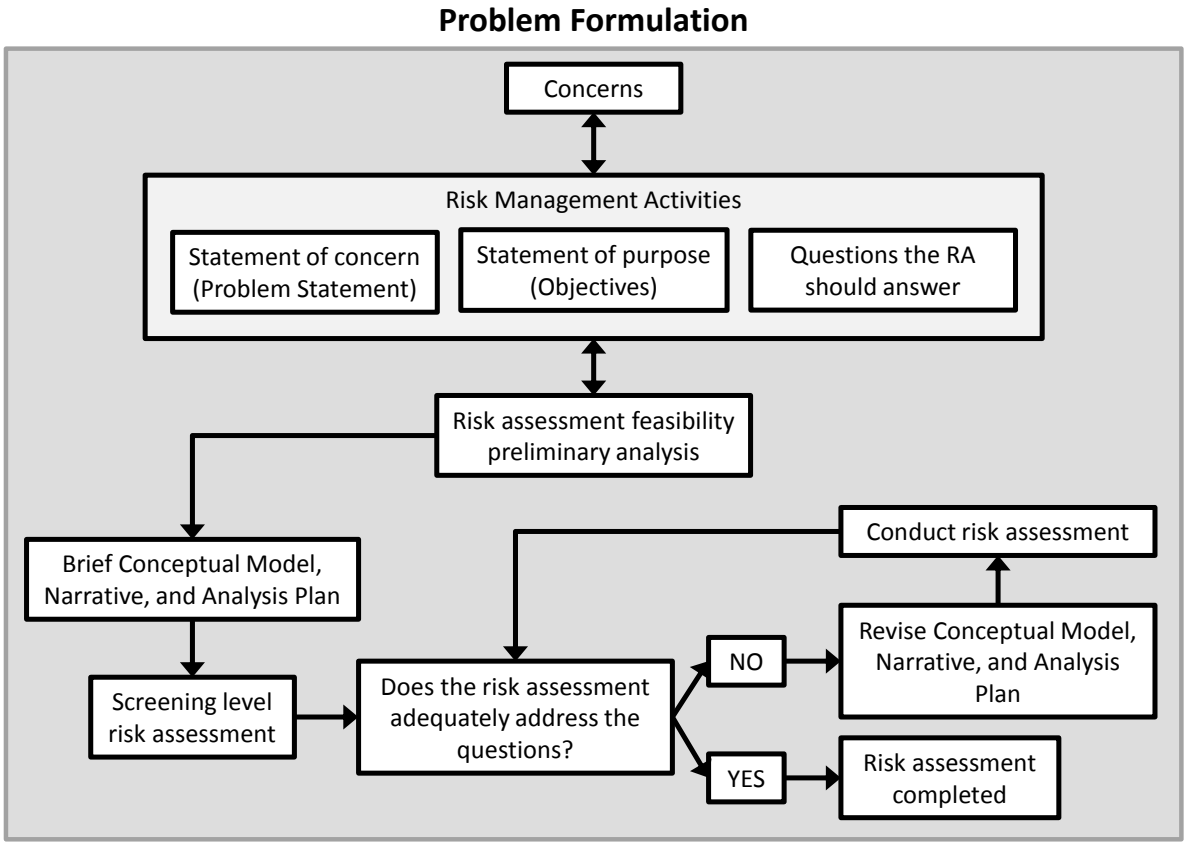


Figure 4. Enhanced Problem Formulation Process Diagram (Source: Adapted from U.S. EPA, 2003c)

It should also be noted that risk assessments can be developed in phases. As indicated previously, a screening level risk assessment may be the initial step that later leads to an enhanced fully quantitative risk assessment. The complexity of the risk assessment may be incrementally increased by adding new models or parameters or by more rigorously characterizing parameter values (e.g., from point estimate values to a statistical distribution, as described in Section 5.3). In many cases, sensitivity analysis can guide prioritization regarding further data gathering or refinement of parameter estimates. The iterative nature of the problem formulation process should allow for further definition and refinement of possible phases of the risk assessment. If multiple versions of the risk assessment are conducted as a result of this iterative process, the choices for each version (also referred to as phase) of the risk assessment should be tracked and documented.

2.2. Overall Problem Formulation and Planning and Scoping

During the problem formulation process, the purpose of the risk assessment is defined through a dialogue between risk assessors, risk managers, risk communicators, and if appropriate—stakeholders. A valuable aspect of the process is documenting the problem formulation development. The value is that it provides a written record of the justification for the decisions regarding the scope, goals, and necessary documentation of the risk assessment. The form of this documentation can vary depending on the needs of the EPA Office conducting the assessment.

The range of acceptable forms for this documentation ranges from a formal and stand-alone problem formulation document to internal notes kept by the project (or work assignment) manager for the EPA Office conducting the assessment. The final risk assessment report should include all of the problem formulation information for the risk assessment iterations that are being published.

The concepts and in many cases the language used during development of the problem formulation documentation can be used in the final risk assessment document. Depending on the form of the problem formulation documentation, the statement of concern and the statement of purpose can be included as part of an executive summary of the risk assessment document. The scope, questions to be addressed, conceptual model, and data not included can be used in the problem formulation chapter of the risk assessment document. Other planning and scoping documentation can be summarized in the problem formulation chapter, a planning and scoping chapter or, if desired, attached as an appendix. The analytical approaches including tools, data inventory, summary of assumptions, and discussion of recommended factors are reiterated as appropriate in the exposure and human health chapters of the risk assessment document. The summary of assumptions is reiterated in the risk characterization chapter (5), which also includes the discussions of variability, uncertainty, and identified gaps.

For comparison, the risk profile approach developed by Codex for microbiological food risk is presented in Text Box 1. Note that in the Codex paradigm, the risk profile is similar to and serves the same purpose as problem formulation described in this MRA Tools document.

2.2.1. Statement of Concern

A concise statement of concern should be developed during problem formulation to convey, in simple terms, what hazard is being addressed and how it is thought to relate to human health for an exposure scenario.

2.2.2. Statement of Purpose and Objectives

The purpose and/or objectives of the risk assessment should be stated in a concise paragraph. Example language for risk assessments performed for the purpose of derivation of Ambient Water Quality Criteria (AWQC) for a specific pathogen is provided below. Note, the designated use and the national scope might be different in other cases.

> This risk assessment is being performed to support the derivation of Ambient Water Quality Criteria (AWQC) for *[pathogen]* under §304(a) of the Clean Water Act (CWA). These will be nationally recommended AWQC for the protection of the *[insert designated use]* designated use. It should not be implied that the AWQC will be protective of other designated uses, such as *[insert designated uses that are excluded]*. As with other §304(a) AWQC, the AWQC for *[pathogen]* are recommended for adoption by states to be used for total maximum daily load (TMDL) determination. States also use AWQC to help assess whether water bodies are threatened or impaired (§305b or §303d CWA) for the specified designated use.

Text Box 1. Microbiological Risk Profile for Food
(Source: Adapted from CAC, 2007)

A risk profile should present, to the extent possible, information on the following:

1. Hazard-food commodity combination(s) of concern:
    - Hazard(s) of concern
    - Description of the food or food product and/or condition of its use with which problems (foodborne illness, trade restrictions) due to this hazard have been associated

Occurrence of the hazard in the food chain

2. Description of the public health problem:
    - Description of the hazard including key attributes that are the focus of its public health impact (e.g., virulence characteristics, thermal resistance, antimicrobial resistance)
    - Characteristics of the disease, including:
        - Susceptible populations
        - Annual incidence rate in humans including, if possible, any differences between age and sex
        - Outcome of exposure
        - Severity of clinical manifestations (e.g., case-fatality rate, rate of hospitalization)
        - Nature and frequency of long-term complications
        - Availability and nature of treatment
        - Percentage of annual cases attributable to foodborne transmission
    - Epidemiology of foodborne disease
        - Etiology of foodborne diseases
        - Characteristics of the foods implicated
        - Food use and handling that influences transmission of the hazard
        - Frequency and characteristics of foodborne sporadic cases
        - Epidemiological data from outbreak investigations
    - Regional, seasonal, and ethnic differences in the incidence of foodborne illness due to the hazard
    - Economic impact or burden of the disease if readily available
        - Medical, hospital costs
        - Working days lost due to illness, etc.

3. Food production, processing, distribution, and consumption:
    - Characteristics of the commodity (commodities) that are involved and that may impact on risk management
    - Description of the farm to table continuum including factors which may impact the microbiological safety of the commodity (i.e., primary production, processing, transport, storage, consumer handling practices)
    - What is currently known about the risk, how it arises with respect to the commodity's production, processing, transport and consumer handling practices, and who it affects
    - Summary of the extent and effectiveness of current risk management practices including food safety production/processing control measures, educational programs, and public health intervention programs (e.g., vaccines)
      Identification of additional risk mitigation strategies that can be used to control the hazard

Text Box 1. Microbiological Risk Profile for Food
(Source: Adapted from CAC, 2007)

(continued)
4. Other risk profile elements:
- The extent of international trade of the food commodity
- Existence of regional/international trade agreements and how they may affect the public health impact with respect to the specific hazard/commodity combination(s)
- Public perceptions of the problem and the risk
- Potential public health and economic consequences of establishing Codex Microbial Risk Management (MRM) Guidance

5. Risk assessment needs and questions for the risk assessors:
- Initial assessments of the need and benefits to be gained from requesting an MRA, and the feasibility that such an assessment can be accomplished within the required time frame
- If a risk assessment is identified as being needed, recommended questions that should be posed to the risk assessor

6. Available information and major knowledge gaps provide, to the extent possible, information on the following:
- Existing national MRAs on the hazard/commodity combination(s)
- Other relevant scientific knowledge and data that would facilitate MRM activities including, if warranted, the conduct of an MRA
- Existing Codex MRM Guidance and related documents (including existing Codes of Hygienic Practice and/or Codes of Practice)
- International and/or national governmental and/or industry codes of hygienic practice and related information (e.g., microbiological criteria) that can be considered in developing a Codex MRM guidance document
- Sources (organizations, individual) of information and scientific expertise that can be used in developing a final Codex MRM Guidance
- Areas where major absences of information exist that could hamper MRM activities including, if warranted, the conduct of an MRA

2.2.3. History and Context within the Agency

Previous risk assessments addressing the same or similar hazards should be summarized to provide context for the current risk assessment. In particular, if previous EPA risk assessments have been conducted, then the relationship between the current and previous risk assessments should be summarized. Relevant information for presenting updated MRAs may include new mandates, policy developments, technical advancements, risk assessment method and tool advancements, and new or enhanced data sets.

2.2.4. Scope

The scope section of problem formulation outlines the scenarios that the risk assessment will address. It is often helpful to list several options for answering the questions listed below. Then, managers and assessors can engage in a dialog to determine which options will be used. The scope should summarize the following:

1. Which infectious disease hazard is being addressed (e.g., pathogen or pathogen strain[s])?

2. Which human populations will be included in the risk assessment (e.g., general population or subpopulations, or geographically defined populations)? Describe which populations are explicitly included in the risk assessment model, which will be accounted for implicitly, and which populations may be excluded by the risk assessment model (e.g., most extreme behaviors).
3. What health outcomes or endpoints are addressed by the risk assessment, including how the health outcome is measured? Clearly defining the health endpoint is important for transparency and also focuses the scope of the risk assessment (e.g., infection, disease symptom/s, mortality).
4. What unit and routes of exposure are relevant and why (magnitude, duration, frequency units)?
5. For risk assessments designed to derive criteria to set "safe" levels of microorganisms, what level of protection (target risk or risk range) will be provided by the criteria, and what is the technical or policy justification for those criteria?
6. What specific exposure scenarios will be modeled? List specific scenarios the risk managers would like to model (varying the inputs), including desired spatial and temporal features.

2.2.4.1. Risk Ranges

Currently, EPA does not have an Agency-wide policy for defining acceptable or tolerable levels of health-based risk associated with pathogenic microorganisms. In fact, there are various regulatory requirements that influence the degree to which MRAs conducted within the Agency are driven by risk ranges. For example, the illness rates associated with the RWQC are 32 and 36 GI illnesses per 1000 primary contact recreators (U.S. EPA, 2012c). The current policy for drinking water standards is to characterize the degree of protection without specific risk-based targets. In this approach the protective ranges have been influenced mainly from feasibility of measurement and application of control technology, taking costs into consideration. Furthermore, semi-quantitative or qualitative MRAs may be necessary under some conditions, and these assessments may still be meaningful for risk management decisions. For example, it may be possible to evaluate the relative degree of protection from fecal contamination in drinking water sources without quantitatively characterizing the risk associated with a specific health endpoint.

Although acceptable risk and target risk are both numeric values that are determined through science policy decisions they are not necessarily always the same. There may be an expectation among some stakeholders that a certain target risk range is acceptable. However, given that different stakeholders may have different ideas about what is acceptable and what is not, it may be misleading to label a risk range as "acceptable." Risk ranges are values that can be estimated empirically from data. However, there may not be clear or convincing information to determine if historically accepted risk ranges are considered acceptable to current stakeholders or not. When risk ranges are used as a driving force or target for MRA conduct within EPA's Office of Water, the risk range is defined along four dimensions, as described below:

1. Risk range is for a specified population (population can be defined in a variety of ways, such as "general," highly exposed, or highly susceptible).
2. Risk range is associated with a defined health endpoint.

3. Risk range covers a defined time span of exposure.
4. Risk range may also be linked to a specific exposure scenario.

Several representative EPA risk range examples employed currently are presented in Text Box 2.

Text Box 2. Information Used to Establish Risk Ranges and Representative Examples of Risk Ranges Currently Employed by U.S. EPA

Pollutant	Specified Population	Health Endpoint	Exposure Duration[a]	Exposure (Designated Use)
Carcinogen	General or children	Cancer	Lifetime	All water uses
Carcinogen	Highly exposed subgroups	Cancer	Lifetime	All water uses
Indicator bacteria (geometric mean)	General	GI illness	Per recreational event	Primary contact recreation
Cryptosporidium (average densities)	General	Cryptosporidiosis	Daily	Treated drinking water consumption

[a] Exposure duration should not be confused with durations relevant for monitoring protocols.

Specific EPA examples:

- Currently, EPA's surface water program has derived AWQC for chemical carcinogens that generally correspond to lifetime excess cancer risk level of 10^{-6} (1 cancer in a million exposed individuals); however, AWQC may correspond to a range from 10^{-7} (1 cancer in 10,000,000 exposed individuals) to 10^{-5} (1 cancer in 100,000 exposed individuals) (U.S. EPA, 2000c).
- EPA's recreational water quality criteria provide a specific level of public health protection. That level of protection is a 30-day geometric mean illness level that is less than or equal to 32 or 36 GI illnesses per thousand recreation events (depending on the fecal indicator bacteria level selected). Although the duration of each exposure event is daily (or shorter), the duration associated with the protection of the criteria is a 30-day period.
- Under EPA's drinking water program (Safe Drinking Water Act), the Long Term 2 Enhanced Surface Water Treatment Rule (LT2) established source water categories (bins) for *Cryptosporidium.* The level of public health protection that is provided by LT2 was driven by a concern for misclassification of binning and cost feasibility for the number of samples that can be monitored. Thus, the ranges of public health protection provided by LT2 are an outcome of this risk management approach rather than a pre-specified target risk range (U.S. EPA, 2003a,b, 2006a).

Note that in these examples the health outcomes are different. Chemical exposures result in an endpoint based on a health effect, whereas microbial exposures can result in infection that can then result in illness. Not all infections result in illnesses. A morbidity factor can be used to convert infections to illness.

2.2.5. Questions to be Addressed in the Risk Assessment

Microbial risk assessments should be scientifically defensible and relevant to regulatory and public health concerns. Therefore, the risk assessment should be framed within the context of Agency policy. The nature and the specifics of the risk management options that need to be evaluated should be developed during problem formulation so that the risk assessment design can address any questions that the risk managers want answered. The questions are important for transparency

and communication between risk managers and risk assessors. Text Box 3 illustrates this point with three examples of questions that risk managers can ask. There may be two types of questions, (1) questions the risk assessment should be able to answer, and (2) questions that the risk assessors need to have answered by the risk managers for appropriate design of the risk assessment. The second point highlights the need for iterative interaction between risk assessors and risk managers.

Text Box 3. Examples of Risk Management Questions that Could Motivate an MRA Investigation

- What effects have broad-based health programs or specific actions (e.g., health education about disinfection) had on (1) the risk of a specific disease (e.g., cryptosporidiosis) and (2) acute gastrointestinal illness risks among children?
- Which pathogens are associated with human health risks from a specified exposure scenario (e.g., freshwater recreation activities)?
- Are there reduced risks to public health associated with implementation of specific water treatment technologies?

2.2.6. Conceptual Model and Narrative

A conceptual model is a graphical representation of the real-world scenario that is being addressed in a given risk assessment (U.S. EPA, 2002b). There should also be an accompanying narrative that explains the conceptual model. The scope of the risk assessment should be consistent with the conceptual model.

The EPA problem formulation workshop (U.S. EPA, 2003c) recommended that multi-tiered conceptual models be constructed. The first (top) tier of the model should be relatively simple, representing only the major components of the assessment. Sub-tier conceptual models can build in more complexity and may require several iterations. The conceptual model should reflect the uniqueness of the situation that is to be addressed. In some cases, a visual diagram that represents how the risk assessment is assembled in the actual software code may serve as a useful sub-tier conceptual model. Although useful for documenting the technical details of the risk assessment, this type of software code map may not clearly communicate the concepts, so should not be solely relied upon as a conceptual model. Collectively, the conceptual model(s) and its narrative should do the following:

- illustrate the risk hypothesis (e.g., provide a flow chart of how risk is thought to occur within the context of the risk assessment scope);
- outline the tools needed to assess the risk (statistical and other models);
- identify available databases that are needed;
- identify default assumptions;
- show what the risk assessment will or will not be able to do, including whether the assessment is quantitative or qualitative;
- summarize data gaps and quality of data;
- consider the interactions between agent, host, and environment when evaluating risk;
- define key uncertainties;

- identify nodes in the risk assessment, including a brief description of the node and what can happen at the node[9]; and
- identify management actions and places where interventions can take place.

Some of the benefits of developing a conceptual model include the following (from U.S. EPA, 1998b):

- The process of creating a conceptual model is a powerful learning tool to inform the conduct of the MRA.
- Conceptual models are easily modified as knowledge increases.
- Together with their narrative description, conceptual models highlight what is known and not known and can be used to plan future work.
- Conceptual models can be a powerful communication tool; they provide an explicit expression of the assumptions and understanding of a system for others to evaluate.
- Conceptual models provide a framework for prediction and are the template for generating more risk hypotheses.

It is important that the conceptual model remain free of risk assessment process elements because trying to reflect the risk assessment process in the conceptual model weakens the conceptual model's ability to represent real-world scenarios. Therefore, the risk assessment processes (allocation of Agency resources and deliverable schedule) should be represented separately from the conceptual model. Because diagrams can be interpreted in different ways by different people, it is essential that a narrative accompany the conceptual model diagram. Details about elements should be included in the text and not clutter the diagram.

Although the concepts of problem formulation and risk assessment can be separated and discussed in a linear manner, the actual process of problem formulation and risk assessment development is an iterative process. The problem formulation stage should be revisited as the risk assessment takes shape. Defining the scope of the risk assessment and choosing an appropriate model may require several iterations, especially if the risk assessment addresses risks or scenarios that have not been modeled previously.

During problem formulation and developing the first drafts of the risk assessment, it should be possible to determine how complex the risk assessment model needs to be to address the questions posed by risk managers. In some cases where the risk assessment questions are simple and limited in scope, a qualitative risk assessment or a simple risk assessment model may be adequate—even when robust data sets are available. As a general guideline, models should only be as complex as they need to be to address the specific risk management questions. A useful model can help the Agency allocate resources and develop a research agenda as well as provide transparency. A simplified model may help the public better understand the process and should thus accompany a very complex model. Within this context, the conceptual model can also be used by the Agency to consider resource allocation and to develop a research agenda.

[9] For example, rainfall, sunlight, and wind speed/direction could be separate nodes in a microbial risk assessment. Relevance of nodes can be evaluated by performing sensitivity analysis.

Figure 5 presents an example overview or top-tier conceptual model for a risk assessment. In this example, the model summarizes how waterborne risk from *Cryptosporidium* is thought to occur. The conceptual model diagram is a visual representation of the risk hypotheses. The risk hypotheses are the proposed answers to risk assessment questions about how exposure occurs and what endpoints are important for the human health hazard. It should be noted that risk hypotheses are not equivalent to statistical testing of null and alternative hypotheses. However, predictions generated from risk hypotheses can be tested in a variety of ways, including standard statistical approaches (U.S. EPA, 1998b). The top tier model should clearly indicate how exposure occurs to provide a conceptual understanding of the magnitude, duration, and frequency of exposure.

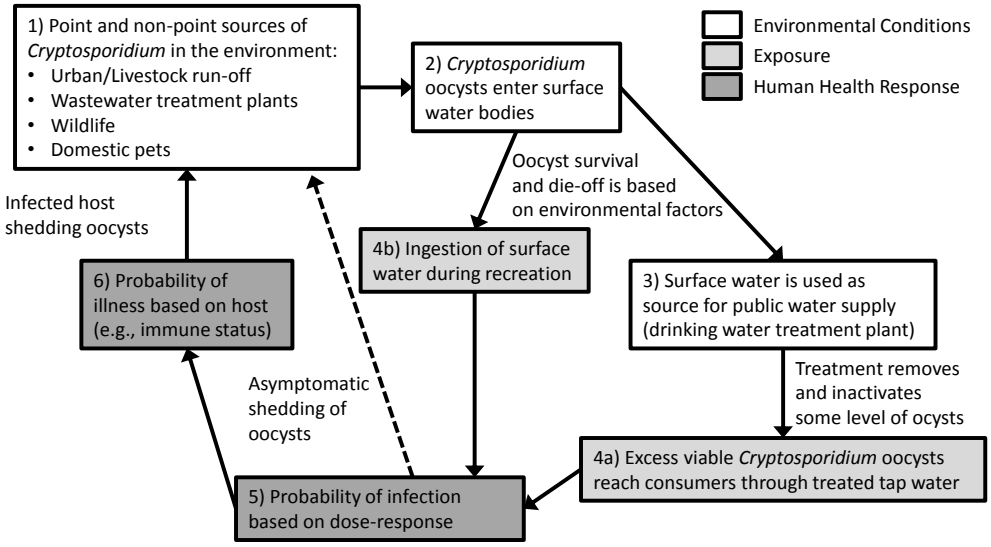


Figure 5. Example of an Overview (Top-Tier) Conceptual Model

2.2.7. Planning and Scoping: Analysis (Operational) Plan

The operational plan should include strategies for dealing with data needs, peer review plans, and any other relevant logistical needs. Information such as lists of relevant experts (for consultation or data contribution) and literature search strategies can be included. This plan may contain a risk assessment process diagram that is a graphical representation of the operational plan that helps explain the logistics of conducting the risk assessment. The plan can also outline proposed phases for the risk assessment as well. Other essential management activities that are part of planning and scoping include timelines, planned deliverables (e.g., status briefing memos, draft for peer review, final draft), team assignments, and possibly budget details. Planning and scoping activities beyond the core scientific issues of problem formulation may be referred to in the risk assessment if those details help increase understanding and transparency.

2.3. Analytical Approaches

For the purposes of this MRA document, the concept of parsimony is encouraged; that is, models should be as simple as possible, but no simpler. Within this context, more complex models should be considered or used under conditions in which the added complexity may provide sufficient

additional insight that the additional complexity is warranted (King et al., 2008; Soller and Eisenberg, 2008). In selecting a MRA model, caution must be taken to ensure that any simplifying assumptions that are employed are in fact appropriate from an epidemiological perspective. Within that context, and to the extent possible, MRAs should use epidemiological data as fundamental components of the assessment and should demand higher quality input data and fewer simplifying assumptions when seeking increased risk assessment accuracy and precision.

The problem formulation should include information on the following topics:

- **Tools.** The tools section of the problem formulation should indicate what software will be used for the risk assessment and may include why the software was chosen. The tools list should also include mathematical tools such as options for dose-response models. "Tools" is also an appropriate section to describe the methods that will be used for dealing with uncertainty. Other types of methodological tools can also be presented.
- **Data Inventory.** The data inventory should list publications that might be consulted during the risk assessment process and sources of data that are being considered for the risk assessment. The list does not need to be comprehensive in the beginning and can be presented in an appendix of the problem formulation if it is overly long. The data inventory may be a work in progress throughout the risk assessment. The data inventory can refer to a literature search strategy that can be presented in an appendix. Literature search strategies should identify which search engines and databases will be used, keywords, key authors, language limitations, and timeframe for the search.
- **Summary of Assumptions.** The summary of assumptions can be organized in different ways; however, listing assumptions that are related to essential risk assessment factors is a systematic way to start. How assumptions limit the scope of the risk assessment and contribute to uncertainty should be explained. The assumptions can be modified and updated as the risk assessment develops.
- **Sources of Variability and Uncertainty.** The sources of variability and uncertainty should be introduced in this section, which should also describe the degree to which variability and uncertainty is or is not captured in the assessment. The iterative nature of problem formulation allows this list to be modified as the risk assessment scope is defined.
- **Factors and Data not Included and Explanation of Why.** There may be information that is not used or avenues not pursued in the risk assessment. The explanation for not including that information should be presented, particularly if other related or similar types of risk assessments have included the information.
- **Identified Gaps in the Knowledge Base.** Although gaps and data limitations may be noted throughout problem formulation, they should also be summarized. Gaps can include a lack of adequate analytical or statistical methods and/or appropriate data and data quality. The summary of knowledge gaps can be useful for prioritizing future resource allocation (e.g., research and development needs) within the context of the results of the risk assessment. Knowledge gaps and data limitations can also affect the number and type of assumptions used in the risk assessment.
- **Environmental Sampling Strategies and Analysis Methods.** Any issues associated with environmental sampling and analysis should be outlined during problem formulation so they can be fully considered during risk characterization. For microbial enumeration, issues may include percent recovery from different sample matrices and the ability of a

method to determine viability. The accuracy, precision, and biases should be included in the description of the methods and protocols.

- **Evaluation of Management Practices.** Depending on the scope of the risk assessment, it may be appropriate to identify which components in the risk assessment can influence or be influenced by management actions. It may be desirable to incorporate scenarios in the risk assessment that include evaluation of best management practices.

2.3.1. Representative Model Forms for MRA Risk Estimation

A variety of model forms can be employed in MRA. Regardless of the form of the model, these models necessarily include exposure and health effects (dose-response) components. Thus, the choices made during the problem formulation phase serve as critical components of the risk assessment. Particular characteristics of each model form allow for the capture of different aspects of the disease transmission system (U.S. EPA, 2004c). In the following sections, several of the most commonly employed models are summarized and reviewed. Exclusion from the following discussion should not preclude use of a particular model form; however, justification for use of a particular model form should be included in the risk description. An overview of two commonly employed classes of MRA models is provided in Table 2, and features of the models that may be used in selecting the appropriate model for a particular application are shown in Figure 6. The model forms summarized in Table 2 (Static and Dynamic) differ in that dynamic models specifically account for the temporally changing effects of person-to-person transmission and immunity in a population, whereas static models treat these innate characteristics as constant modulators of population risk.

Table 2. Overview and Comparison of Static and Dynamic Risk Assessment Models

Static Risk Assessment Model	Dynamic Risk Assessment Model
Number of susceptible individuals is time invariant	Number of susceptible individuals varies over time
Environment-to-person	Environment-to-person, person-to-person, and person-to-environment-to-person
Individual-based perspective	Population-based perspective
Typically assumes that the potential for secondary transmission of infection or disease is negligible or scales linearly with the number of infections	Typically account for the potential for secondary or person-to-person transmission of infection or disease
Typically assumes that immunity to infection from microbial agents is negligible	Exposed individuals may not be susceptible to infection or disease because they may be infected already or may be immune from infection due to prior exposure
Dose-response function is the critical component in a quantitative risk assessment	The dose-response function is important; however, person-to-person transmission and immunity may also be important

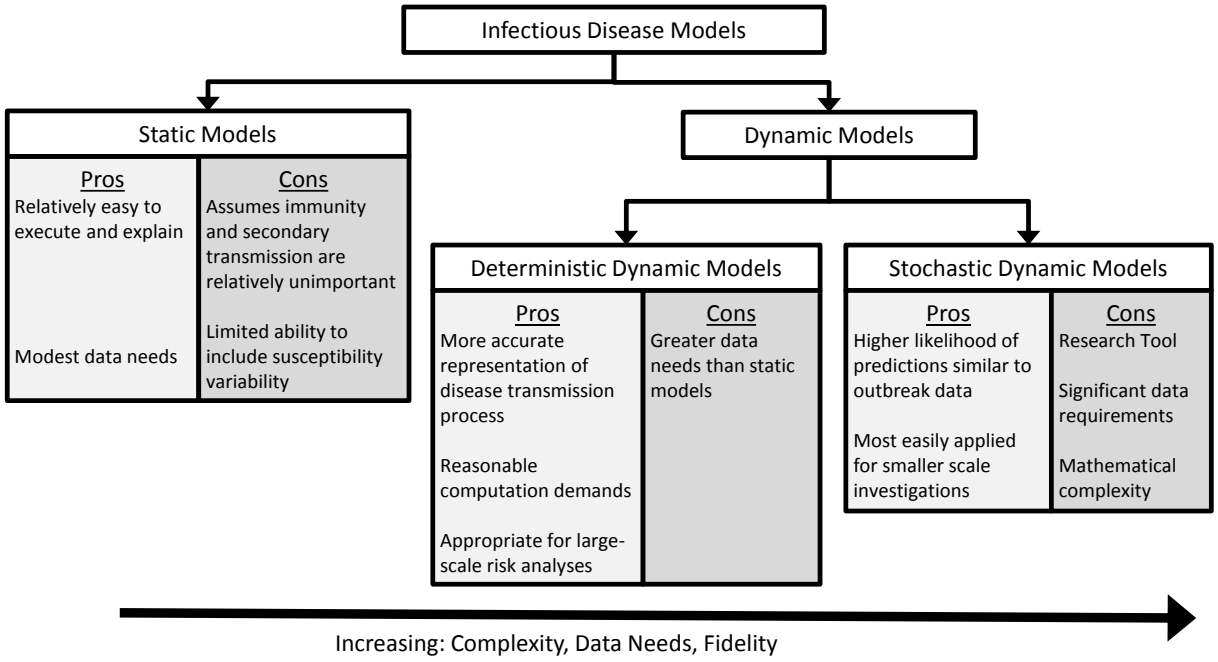


Figure 6. Infectious Disease Model Features for Use in Model Selection

2.3.1.1. Static Models

Some infectious diseases are not readily transmitted from person-to-person but are acquired, to the best of current knowledge, only by consumption of or contact with contaminated environmental materials (e.g., *Mycobacterium avium* complex [MAC] infection from drinking water). In other cases, although an agent may have the potential to be transmissible, the person-to-person component is unknown or thought to be negligible.

Understanding the pattern of human infections from such pathogens or exposure scenarios may be best achieved through the use of static models (parallel to those used for toxicological risk assessments). The chemical risk assessment-based models are used to estimate risk at an individual level and typically focus on estimating the probability of infection or disease to an individual as a result of a single exposure event. With respect to microbial contaminants in water, a fundamental simplifying assumption of static model-based analysis is that exposure events and infection/disease are independent; that is, the outcome from one exposure event does not affect a subsequent exposure, and one individual's outcome has no effect on any other individual's outcome. In most static models, it is assumed that the population may be categorized into two epidemiological states—a susceptible state and an infected or diseased state. In these models susceptible individuals are exposed to the pathogen of interest and move into the infected/diseased state with a probability that is governed by the dose of pathogen to which they are exposed and the infectivity (dose-response relationship) of the pathogen.

A static model is appropriate in cases where the probability of infection or illness is not likely to be substantially impacted by population-level factors such as person-to-person transmission. Such models can handle complex details about the course of events that lead to exposure and infection

and can be analyzed by well-established statistical techniques that require fewer assumptions than do dynamic models (discussed below). Static models are useful for analyzing situations where the effect of an intervention directed to individuals (e.g., point-of-use remediation) is more important than the effect on transmission throughout the population; they are not appropriate for measuring indirect effects at the population level (e.g., the effect of water treatment interventions on risk due to secondary transmission).

A representative conceptual model for a static MRA model is presented in Figure 7. As can be seen, individuals who are exposed to pathogens from a specific source, move from a susceptible state into an infected or diseased state with some probability that is governed by their exposure and the dose-response relationship for that pathogen. Also note that previous exposures to the pathogen, interactions with other (potentially infected) individuals, other routes of exposure, and immune status are not included in this type of model. However, it is possible to use these models to estimate the cumulative risk of recurring exposures, provided that those recurring exposures are assumed to be independent (one such example is an estimated annual risk from daily ingestion of drinking water).

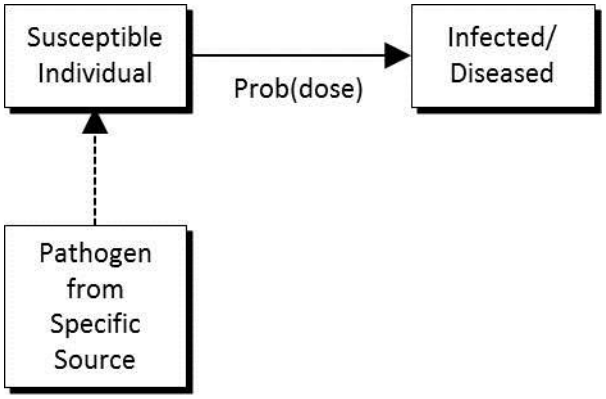


Figure 7. Static Risk Assessment Conceptual Model

2.3.1.2. Dynamic Models

Risk managers and regulators are often concerned with risk on a societal or population scale. For a thorough evaluation of risks that are manifest at the population level, MRA methods must explore the relative importance of secondary transmission and immunity, and thus capture and integrate the dynamic interplay of hosts, agents, and environments.

Secondary cases (often represented in epidemiological studies by a secondary attack rate) generally refer to cases that occur among contacts, within the incubation period of the pathogen, and following exposure to a primary case. In some cases, direct person-to-person transmission cannot be distinguished from contamination of the immediate environment (e.g., toddlers sharing toys versus direct physical contact during play). Depending on the purpose of the assessment, it may be appropriate that the definition of secondary transmission include infections that result from propagation of the specific exposure of interest, but not encompass distant transmissions (separated by time and/or space) that may be more appropriately considered to result as a function of person-to-environment-to-person transmission. Temporal and spatial limitations can be specifically noted in the definition of secondary transmission.

MRA models can be configured to account for secondary transmission and immunity in a population through the use of a dynamic model (Anderson and May, 1991). These models, which can take several forms (deterministic or stochastic), characterize the dynamic epidemiological status of the population (e.g., susceptible to infection, symptomatic infection, immunity). Static MRA models do not account for the dynamic nature of secondary transmission, although dose-response parameters derived from static models may be incorporated into dynamic models. Inclusion of secondary transmission in MRAs can provide non-intuitive results (Eisenberg et al., 2008); therefore, if secondary transmission and other innate characteristics of infectious disease transmission are not included in the assessment, a sound justification for this decision is a suggested component of the risk assessment documentation.

The use of these transmission models in MRA has increased in the past 10 -15 years and there are numerous examples of such models in the literature. For example, Zelner et al. (2010) use a transmission model to examine secondary spread through households after a point source foodborne outbreak. Eisenberg et al. (2005) used transmission models to analyze the 1993 *Cryptosporidium* drinking water outbreak focusing on (1) disaggregating the risk associated with direct exposure to the contaminated water and subsequent secondary spread; (2) assessing the role that person-to-environment-to-person played in the outbreak, and (3) assessing the role that immunity played in the outbreak. Sheng et al. (2009) provides a framework for examining environmental infection transmission systems and Eisenberg et al. (2002) provides a policy perspective for using transmission models in decision-making.

Dynamic MRA models take two main forms—deterministic and stochastic. "Deterministic" means that the model output is strictly determined by the starting conditions and the values of the parameters in the equations that define the system (i.e. point estimates). In stochastic models, events are treated as stochastic (random) events within a distribution rather than deterministic ones. Dynamic MRA methods have been used for numerous specific case studies in the United States (Eisenberg et al., 1996, 1998; Koopman et al., 2002; Soller et al., 1999, 2003, 2006) and recently to support regulatory decisions by EPA (U.S. EPA, 2006b). Stochastic MRA models are still research tools that continue to develop.

Deterministic Dynamic MRA Models

Deterministic dynamic MRA models are suitable for large populations of individuals randomly interacting with one another. In this form, the population is divided into one of the following different epidemiological states: (1) susceptible, (2) diseased (infectious and symptomatic), (3) carrier (infected but asymptomatic), and (4) immune (partial or complete). Only a portion of the population is in a susceptible state at any point in time, and only those individuals in a susceptible state can become infected through exposure to pathogens. The dynamic aspect of the model means that members of the study population move between epidemiological states at different rates, and thus, the number of individuals in each state changes over time.

Variables in the model track the number of individuals that are in each of the epidemiological states at any given point in time (thus, these variables are called state variables). The sum of the number of individuals in each of the epidemiological states equals the total population. A representative conceptual model for this type of MRA model is presented in Figure 8.

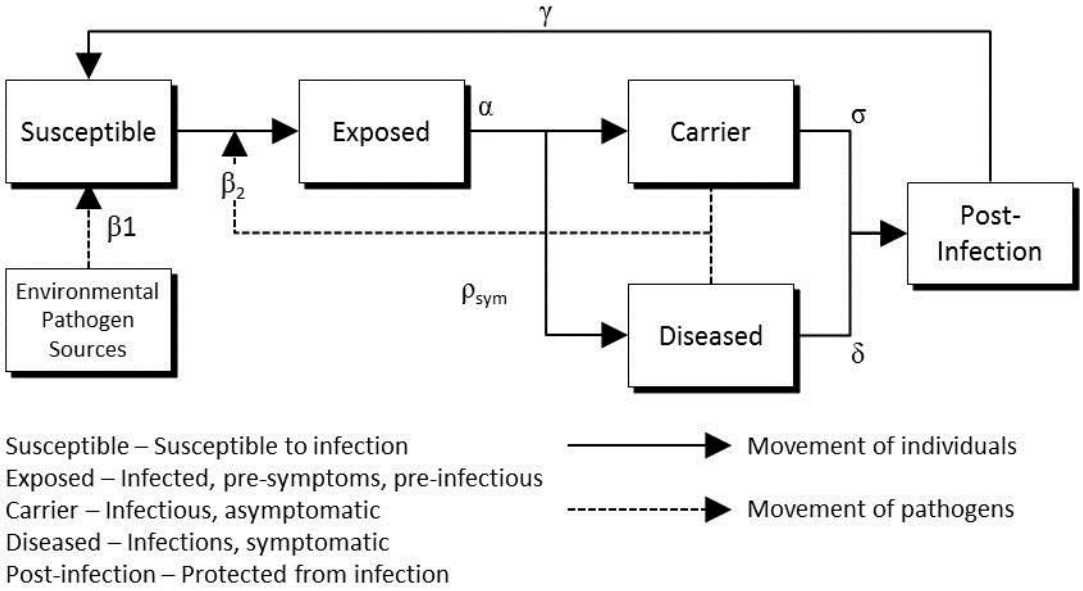


Figure 8. Dynamic Risk Assessment Conceptual Model (Source: Soller and Eisenberg, 2008)

Deterministic dynamic MRA models are expressed mathematically as a set of differential equations. These equations describe the rate of change in the number (or density) of individuals in a particular state (or compartment) over time and have defined parameters and starting conditions. Rate parameters (i.e., the Greek letters in Figure 8) determine the population's movement from one state to another. Factors affecting the population dynamics include the level and frequency of exposure, the ability of individuals in infectious states to infect susceptible individuals, and the temporal processes of the disease (e.g., incubation period, duration of disease, duration of protective immunity). The rate parameters may be determined through literature review or through site-specific data, if available and appropriate. Whether single or multiple exposures are considered should also be discussed during the problem formulation stage.

Deterministic dynamic MRA models have a number of limitations. Modeling relatively small populations can lead to misestimation of disease. These models also require appropriate parameter values for transmission rates, and such information can be difficult to determine accurately. Lack of knowledge and data as well as inherent biological variability suggest a need for uncertainty and sensitivity analyses of parameter values.

Finally, comparison of static and deterministic dynamic models indicates that under a specific set of assumptions, the two models are essentially equivalent (Soller and Eisenberg, 2008). The conditions in which a static model would yield similar results to a deterministic dynamic model are as follows:

- the background density of the pathogen (or equivalently, the endemic level of infection/disease) in the population is zero or unimportant;

- the duration of infection and disease approaches zero; and
- infection and/or disease do not confer immunity, or the duration of immunity approaches zero.

Stochastic Dynamic MRA Models

In a stochastic form, dynamic models incorporate probabilities at an individual level and are evaluated by an iterative process (e.g., susceptible person A has a probability of contacting person B, who has a probability of being infectious). This type of model also uses states (or compartments) for classifying the epidemiological status of the population and subpopulations (e.g., HIV-positive individuals, individuals greater or less than 5 years of age) under study, but differs from the deterministic dynamic MRA models in that the compartments contain discrete individuals rather than the numbers or densities of persons.

In stochastic dynamic MRA models, events are treated as random (stochastic) events within a distribution rather than deterministic ones. These models employ distributions of outcomes rather than an average of outcomes as used in deterministic models. A stochastic model will produce different results, within a range, each time it is run—even with the same starting conditions and parameters due to the effects of chance. Stochastic forms are suitable for small populations and heterogeneous mixing. In a small population, chance events, such as an infectious person contacting only immune persons during the infectious period of illness, can have a substantial effect on the transmission dynamics of the disease (U.S. EPA, 2004c).

These types of models have been used to investigate the stochastic effects of disease transmission and localized exposure (U.S. EPA, 2004c). For example, King et al. (2008) used a nonlinear stochastic model coupled with a new likelihood maximization procedure for model parameter values to explain the dynamics of cholera infection in Bengal, the pathogen's endemic home.

2.3.2. Data Representation in MRA Risk Estimation Models

In assessing risk associated with infectious disease hazard exposures, it is necessary to estimate a number of parameters in the risk models. Depending on the data quality, different representations of these data (as discussed below) may be appropriate. For some chemical risk assessments, EPA has made a policy decision that conservative (more protective) estimates of some exposure factors should be used to assure the desired level of health protection for sensitive segments of the exposed population (U.S. EPA, 2000c). The Agency has not developed a comprehensive policy regarding how conservative parameter estimates should be in MRA. In fact, the use of multiple layers of conservative estimates for microbial contaminants has been shown to result in risk estimates that are not credible and that are overly protective (U.S. EPA, 1995a). Thus, the selection of values used in the risk assessment and the respective data representation should be well documented in the risk description. The following is an overview of the various ways that data can be represented in an MRA.

2.3.2.1. Point Estimates

A "point estimate" is a single-valued estimate of a parameter used in risk assessment. Using point estimates for all the parameters in a risk equation results in a single value (point estimate) of risk that provides no information concerning the potential sources of variability or uncertainty or the magnitude of that uncertainty associated with the risk estimate. Lack of information regarding potential variability and uncertainty in the quantity being estimated is a fundamental weakness of the point estimate approach. The strength of using point estimates is the relative ease of use and simplified risk assessment output. In some cases, the point estimates themselves may be selected taking the potential uncertainties in the parameter values into consideration. For example, EPA's *Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health* (U.S. EPA, 2000c) recommends that 90th percentile estimates derived from national studies of drinking water and fish consumption by adults be used in estimating criteria values for chemicals.

Confidence limits provide an indication of the degree of uncertainty associated with a statistic (Snedecor and Cochran, 1989). Although they are usually derived for estimates of the arithmetic mean, they can also be estimated for other statistics (e.g., median and percentiles). The narrower the interval, the more precisely the statistic has been estimated. The magnitude of uncertainty is expressed in the form of upper and lower confidence limits (collectively known as the confidence interval); confidence limits always have an associated confidence level (e.g., 90%, 95%). The confidence level reflects the estimated probability that the numeric statistic estimated, based on a sample of a given population size, will fall within the specified confidence interval. Confidence limits typically assume that the underlying distribution in the study population is "normal" (Gaussian), but alternative assumptions can also be used. Confidence limits can also be derived that make no assumptions (nonparametric) about distribution shape. An example of a confidence limit can be found in EPA's *Risk Assessment Guidance for Superfund* (U.S. EPA, 1989), where the 95% upper confidence limit on the arithmetic mean soil density is recommended as the appropriate point estimate for a screening level risk assessment at Superfund sites.[10]

2.3.2.2. Statistical Distributions

If adequate data are available, it may be possible to accurately characterize the statistical distribution of a parameter used in risk assessment. That is, there may be enough data to select the form of the distribution and to accurately estimate its parameters (e.g., mean, standard deviation, percentile values for a normal or lognormal distribution). Where such data are available (examples include national surveys of water intake and body weight), individual summary statistics can be estimated very accurately (i.e., confidence limits are narrow).

2.3.2.3. Bayesian Methods

The same methods that are used in dose-response modeling can also be used to characterize the uncertainty in model parameters through the generation of "uncertainty samples." These uncertainty samples are particularly useful in MRA because they fully characterize the uncertainty for a specific model parameter, given the available data. For example, Messner et al. (2001) combined three *Cryptosporidium* isolates that were considered representative of a larger population of human-infecting strains and determined that the risks of infection produced from

[10] For up to date information on Superfund risk assessments visit: http://www.epa.gov/oswer/riskassessment/risk_superfund.htm

single oocyst doses for a mixture of the three isolates and for an oocyst selected at random from the larger population of strains were 0.018 and 0.028, respectively. A related uncertainty analyses was conducted for the MRA that supported the economic analysis of the Long Term 2 Enhanced Surface Water Treatment Rule (LT2) drinking water regulation (U.S. EPA, 2006a).

Hierarchical Bayesian modeling is often used in MRAs to combine results from different studies or isolates in a meta-analysis. For each study or isolate, the parameters can be randomly selected from a distribution that depends on several additional parameters, called hyperparameters. For example, the single oocyst infection probability for each *Cryptosporidium* isolate can be modeled as being randomly drawn from a normal distribution with hyperparameters μ and σ that represent the variability of the isolate infectivity across the isolate population (Messner et al., 2001). Markov Chain Monte Carlo method (MCMC) methods can accommodate hierarchical Bayesian models (Section 4.2.1 and Appendix C).

For example, Gronewold et al. (2009) demonstrated that Bayesian techniques can be used for quantifying and analyzing uncertainty in exposure model fate and transport parameters. In that study, Bayesian methods for addressing uncertainty and developing models were compared with regression techniques in which a model was assumed and uncertainty was assumed related to confidence in the estimates of model parameters. In comparing approaches for estimating decay rate parameters from microbial survival experiments, Gronewold et al. (2009) found that Bayesian techniques, because they rely on fewer assumptions about parameter variability than alternative techniques, provided higher estimates of variability in the parameters and likely reflect actual conditions more accurately. Bayesian techniques also allowed these researchers to assess the forms of models proposed for microbial inactivation and to assess alternative models of the process. The work reported in Gronewold et al. (2009) is an extension of prior studies by the authors (Gronewold et al., 2008) in which uncertainty in different enumeration processes was quantified and related to assessment of water quality.

EPA anticipates that hierarchical modeling will be important in the future of microbial risk assessment. Roles that Bayesian techniques may be expected to play include development of dose-response models in the absence of human dose-response data, parameter estimation for sparse data sets or for data sets exhibiting wide variability, or assessment of alternative models, particularly in exposure assessment. Bayesian methods are further discussed in Section 4.2.1 with respect to their use in dose-response modeling and are discussed more fully in Appendix C.

2.3.2.4. Probabilistic Simulations

Distributional data and/or Bayesian-based uncertainty samples can be used in probabilistic MRAs. In these types of analyses, risk calculations (each of which yields a point estimate) are repeated many times (typically thousands of times) using random or structured "draws" of values from the distributions of each parameter value. The resulting distribution of risk provides information about the expected precision of the estimate, given the distributions of and/or uncertainty associated with the input parameters. The contributions of variability in individual parameters can also be estimated and the correlations among parameters can be accommodated within a Monte Carlo framework, described below.

EPA has developed guidelines for when probabilistic methods can and should be used in health risk assessments (U.S. EPA, 1997b). The most common obstacle to the use of probabilistic modeling is the lack of data to adequately characterize the variability and/or uncertainty in key input parameters. One approach that has been used at EPA is a "tiered approach" to risk assessment, whereby the first step is a set of screening calculations to determine if the risks being estimated fall within the range of concern under a credible set of assumptions. If the results of the screening level analysis warrant further evaluation, sensitivity analyses can be used to further characterize the likely range of risks and to guide data gathering efforts for key parameters. If sufficient data are available, and if more detailed information is needed or desired regarding the decision being evaluated (e.g., setting a health-based criterion), then Monte Carlo modeling may be useful as a subsequent tier.

A Monte Carlo simulation is a statistical technique for evaluating the range of possible outcomes of multiple processes whose outcomes or inputs are random variables. The alternative to Monte Carlo simulation—integration over the distribution of possible values for the random variables—is often mathematically impossible. Additionally, many software packages or programming languages make even large number of Monte Carlo simulations an easy operation on personal computers. In MRA Monte Carlo simulations, random values for the variables in a mathematical model for estimating risk are drawn from appropriate distributions. For each simulation, the risk is calculated based on the mathematical model using the values drawn from their respective distributions. The results of many simulations may be combined as a distribution of risks associated with the system. This distribution of risks allows more comprehensive characterization of risk than a simple point estimate. For example, risk management decisions might be made based on the 95th percentile of the distribution resulting from the Monte Carlo simulation. The distribution of outcomes permits a more informed selection of risk-based decisions considering the full range of data inputs and the nature of the effects of concern. It can also support consideration of costs and benefits arising from key decisions. Use of Monte Carlo simulations requires significantly more information about the system being modeled than generation of point estimates. Each random variable is characterized by a distribution (either parametric or non-parametric) and the choice of the distributional form and estimation of distribution parameters both require data or some other information that informs their selection.

Several illustrations of the Monte Carlo modeling technique within QMRAs of drinking water exposures are provided below to show the importance of considering variability in processes within a MRA and to illustrate the application of the technique. Olivieri et al. (1999) used Monte Carlo simulations of alternative advanced water treatment (e.g., reverse osmosis) trains to compare the efficiency of the entire treatment trains with respect to removal and inactivation of an enteric virus surrogate. The data used in that study were generated in pilot studies of the unit operations. The Monte Carlo simulation allowed exploration of interactions in performance of individual unit operations prior to building a full-scale facility. Distributions for removal in each of the unit operations were based on observed removal from challenge studies of pilot processes and are presented in Table 3. In this study Monte Carlo techniques provided an opportunity for evaluation of multi-component treatment processes and for providing quantitative information for processes expected to reduce microorganism density below detectable limits.

Table 3. Distributions used in Monte Carlo simulations conducted by Olivieri et al. (1999)

Unit Process	Distribution used to characterize removal of viruses	Justification
Influent	Log-normal	Maximum likelihood estimation
Reverse osmosis filtration	Weibull and gamma	Maximum likelihood estimation
Ultrafiltration	Log-normal	Insufficient data for maximum likelihood estimation; distribution assumed log-normal
Ozonation	Point estimate	Insufficient data to characterize distribution
Microfiltration	Gamma	Maximum likelihood estimation
Chlorine contactor	Uniform distribution	Based on CT (contact time) operational range

Pouillot et al. (2004) used a second-order Monte Carlo simulation to estimate cryptosporidiosis risk based on observed *Cryptosporidium* densities in a finished water reservoir and to initiate discussion of monitoring strategies and alternative water quality standards for managing risk. This study is reviewed here because of its separation of uncertainty and variability via the second order Monte Carlo simulation. It is also notable in its explicit modeling of both immune-competent and immune-compromised individuals. The second order Monte Carlo simulation entailed using a Monte Carlo simulation to develop a distribution of a parameter making up the system model, then sampling from the resulting distribution in a Monte Carlo simulation of the overall system. In this case, the inner Monte Carlo simulation developed distributions of uncertain parameters and the outer Monte Carlo sampled from those distributions systematically to develop a distribution for risk. Other studies (Petterson et al., 2007; Signor et al., 2005) have noted the importance for explicitly modeling parameter uncertainty within MRAs and proposed the use of second Monte Carlo simulations in such analyses.

The decision whether to use probabilistic methods can be technically complex; thus, expert statistical advice should be sought to support such decisions. When planning such assessments, it is important to ensure that the approach taken to characterize uncertainty is consistent across the models used in all stages of the risk assessment. An example of such an analysis can be found in EPA's risk assessment in support of the LT2, which addresses *Cryptosporidium* contamination in sources of drinking water (U.S. EPA, 2006a).

2.4. Elements to Consider During Problem Formulation

The problem formulation process should provide a working outline of the risk assessment. Furthermore, information that is required for the exposure assessment and health assessment phases of risk assessment is preliminarily gathered and reviewed during problem formulation. The elements also have the potential to influence one another. Where appropriate, these influences should be noted during problem formulation and included in the risk assessment documentation.

In this document, the infectious disease hazard characterization and host characterization are initially considered as part of problem formulation because the resulting data and information are important for building the conceptual model(s) and making the decision if adequate data are available for the desired scope of a given risk assessment. It is appropriate to consider these steps as overlapping with the risk assessment phase because the data gathered during problem formulation are then used during the risk assessment.

2.4.1. Infectious Disease Hazard Characterization

For the purposes of this document, an infectious disease hazard is defined as a pathogenic microorganism. If you wish to consider a surrogate for an infectious disease hazard, such as fecal indicator bacteria, please refer to EPA's Technical Support Materials (U.S. EPA, 2014). Infectious disease hazards can also include multiple pathogens simultaneously, such as reported by Westrell et al. (2003), where risks due to failures in drinking water treatment systems were modeled for the pathogens *Cryptosporidium*, rotavirus, and *Campylobacter jejuni*.

Factors related to infectious disease hazards that should be considered during problem formulation are listed and briefly discussed below (also see Table 1):

Infectious disease hazard characterization elements include (adapted from ILSI, 2000):

- survival, multiplication, and accumulation;
- resistance to control or treatment processes; and
- ecology (including zoonotic potential, vectors, and epidemiological triangle).

Pathogen elements that overlap between exposure and human health effects include:

- virulence and pathogenicity of microorganism;
- pathologic characteristics/disease caused, including host specificity (including zoonotic potential and vectors);
- infection mechanisms/route of infection/portals of entry;
- potential for secondary transmission; and
- taxonomy/strain variation.

Environmental Survival, Multiplication, and Accumulation

A pathogen may be able to survive in water but be unable to infect a host. Many molecular-based microbial assays and some fluorescent antibody assays do not distinguish between live/dead or

infectious/noninfectious organisms (e.g., deoxyribonucleic acid amplification methods, polymerase chain reaction [PCR]). For assays that require growth of the microorganisms under laboratory conditions there is a concern that VBNC microorganisms will not be detected. Risk assessors should be aware of, and report the caveats of, the assays used to quantify microorganisms in the studies they use as data sources for an MRA.

Multiplication refers to the ability of some microorganisms to reproduce or grow in the environment. The combination of survival, viability, infectivity, virulence, and multiplication may be addressed through fate and transport modeling. Accumulation can occur in a variety of ways. Some examples include accumulation in biofilms (in pipes or tanks), accumulation in sediments, adsorption to particulate matter in water, and bioaccumulation in filter feeding aquatic organisms (e.g., shellfish). These factors contribute to heterogeneity of microbes. Places in the risk scenario where accumulation can occur should be noted.

Survival, multiplication, and accumulation of microorganisms are dependent on environmental conditions such as temperature, nutrient availability, and other water quality parameters (NRC, 2004). Treatment processes can also influence survival and may alter virulence and pathogenicity. Table 4 presents several representative tools for modeling pathogen survival, multiplication, and accumulation. Environmental niches that can harbor pathogens should be considered, such as biofilms and amoebae (e.g., *Legionella* can live inside amoebae; Brown and Barker, 1999). The extent to which survival, multiplication, and accumulation will affect the risk assessment should be considered and documented during problem formulation.

Table 4. Representative Tools for Modeling Pathogen Survival, and Multiplication

Tools	Reference
Survival and Transport of Viruses in the Subsurface: An Environmental Handbook. This issue paper discusses some of the conditions under which viral contaminants may survive and be transported in the subsurface, identifies sources as well as indicators of viral contamination, outlines the effects of hydrogeologic settings on viral movement, and introduces the reader to the current state of virus transport modeling along with an example of modeling applications.	U.S. EPA, 2003d
Continuous simulation	Recommended by TMDL Protocol (U.S. EPA, 2001)
Monte Carlo simulation	Recommended by TMDL Protocol (U.S. EPA, 2001)
Log-normal probability modeling	Recommended by TMDL Protocol (U.S. EPA, 2001)
USDA/Agricultural Research Service (ARS) Pathogen Modeling Program (PMP) estimates the effects of multiple variables on the growth or survival of foodborne pathogens	Version 7.0 http://www.arserrc.gov/mfs/Download.htm

Table 4. Representative Tools for Modeling Pathogen Survival, and Multiplication (continued)

Tools	Reference
ComBase, also developed by USDA/ARS, is an on-line database of predictive microbiology information collected from researchers, institutions, and published literature. ComBase may be searched based on temperature, pH, water activity, condition, source (publication), organism, and environment. Files are provided giving organism, maximum rate, doubling time or D-value, source, conditions, environment, temperature, pH, water activity, a table and chart for log density versus time, and other available details. (Maximum rate is the maximum slope of the "log [cell density] versus time" curve, in a given environment.)	http://wyndmoor.arserrc.gov/combase/default.aspx

Resistance to Control or Treatment Processes

Microorganisms have varying degrees of resistance to water treatment and control processes. The extent to which these control or treatment processes will affect the risk assessment should be considered and documented during problem formulation. For example, data on how pathogens respond to both wastewater treatment and public water supply treatment should be noted, as appropriate. If the risk assessment is for a performance target, then the treatment and control processes may be of central importance. For example, *Cryptosporidium* oocysts are very resistant to conventional disinfection with chlorine, so chlorination in the absence of filtration may be inadequate to protect public health if oocysts are present in source waters for drinking water.

Ecology

The epidemiological triangle (epi triad) is a recommended model for conceptualizing agent-host-environment interactions and is a useful way to consider ecology (Figure 9). The epi triad can be used to predict epidemiological outcomes and provides a tool to discuss parameters that influence public health outcomes. The epi triad can capture how pathogen, host, and environment all affect each other.

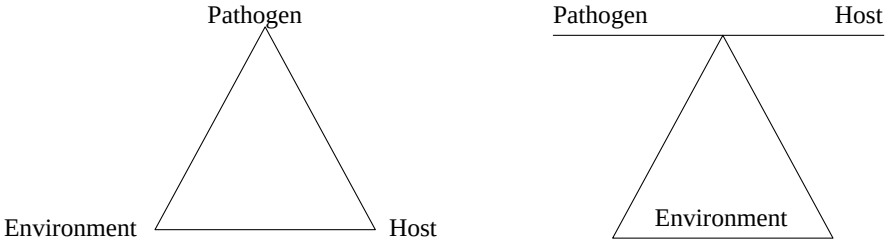


Figure 9. Two Versions of the Epi Triad (Source: CDC, 1992)

Physical properties of microorganisms that relate to their transport/mobility (e.g., hydrophobicity) and data on pathogen survival and bacterial colonization under varying ecological conditions (e.g., stressors such as pH, nutrient availability, and temperature) should be considered and discussed within the context of the scope of the risk assessment. Often, the ecology of pathogens can be

elucidated by examining their transport, fate, and survival in the environment—particularly how they react in variable media. If sufficient data or information is available to prepare an ecology summary for the pathogen, it should focus on the appropriate exposure source[11] (water, food, other) and may include environments that the microorganisms encounter before they enter the media of concern. For example, MAC thrive in hot water and have been known to colonize hot water systems in hospitals and buildings (Primm et al., 2004).

Ecological niches may also be provided by other microorganisms. For example, biofilms create ecological niches that are important to consider because microbes often exhibit different properties in communities compared to the same species living in suspension. For example, *V. cholerae* from human stool has enhanced infectivity in a rabbit model relative to infectivity of dispersed (planktonic) cells (Faruque et al., 2006). This is because *V. cholera*e from human stool are in conditionally viable forms (biofilms and multicellular clumps).

Protozoa can also harbor bacteria within their cell membranes, thereby protecting the bacteria from many environmental stresses. *V. cholerae* occur commensally in zooplankton. A single copepod can carry up to 10^4 cells of *V. cholerae*, and human volunteer studies show that ~10^4 to 10^6 *V. cholerae* can cause clinical cholera (Colwell et al., 2003). *Legionella* are known to reside within at least 20 species of amoebae, two species of ciliated protozoa, and one species of slime mold (Lau and Ashbolt, 2009). Species from the pathogens *Vibrio*, *Mycobacterium*, *Helicobacter*, *Afipia*, *Bosea*, *Pseudomonas*, and mimiviruses are also associated with protozoa in the environment.

Virulence, Pathogenicity, Pathological Characteristics, and Host Specificity

Virulence and pathogenicity refer to how easily and effectively a pathogen may cause disease in a host. Virulence is "the degree of intensity of the disease produced by a microorganism as indicated by its ability to invade the tissues of a host and the ensuing severity of illness." Pathogenicity is "the property of an organism that determines the extent to which overt disease is produced in an infected population, or the power of an organism to produce disease. It is also used to describe comparable properties of toxic chemicals. Pathogenicity of infectious agents is measured by the ratio of the number of persons developing clinical illness to the number exposed to infection" (U.S. EPA, 2007a). Although both can be expressed numerically, the general definitions can be broader. Pathological characteristics are a description of the disease symptoms that result from exposure and infection by the pathogen (including strain variations). The known range of disease symptoms should be briefly reviewed and the specific health endpoint that the risk assessment addresses should be presented within the context of the broader range of health endpoints.

Host specificity is a pathogen characteristic that is related to host susceptibility. A species is not considered a host if it cannot be infected by the pathogen. Note that a species can still be considered a host even if no illness results from infection. Within a host species there is variability in susceptibility. For example, mild illness can occur in immune-competent persons resulting from

[11] Exposure sources can be the media through which the contaminant is delivered, such as water or shellfish, or exposure sources can indicate the origin of the contaminant, such as point source, non-point sources, or naturally occurring. Exposure route indicates the sites of body contact that are relevant for access to sensitive tissues and organs, such as ingestion, inhalation, and dermal exposures.

exposure to a pathogen; whereas, severe illness may occur in immunocompromised persons. However, host specificity most often refers to the range of species that are infected by the pathogen.

Information that can facilitate the comparison of human response to a pathogen versus laboratory animal models' response should be examined and is particularly important if data from animal models will be used to characterize dose-response or symptomatology during the risk assessment. Wild and domestic animals may also be prone to infection and disease (zoonotic potential) and thus may be a source of pathogens for human exposure either directly or through transport in the environment. Some pathogens have non-human carriers, also known as vectors, which are important in the pathogen life cycle or serve as an environmental reservoir.[12] The potential role of susceptible animals, vectors, and environmental reservoirs in the risk scenario should be addressed, which may include an explanation of how animals are contaminating the water sources of concern. These factors are also evaluated in greater detail in the health effects section of the risk assessment.

Infection Mechanisms, Route of Infection, and Portals of Entry

Infection mechanisms, route of infection, and portals of entry emphasize the manner in which pathogens interact with hosts. The exposure routes[13] that will be included in the risk assessment are defined during problem formulation. Part of that definition should include identification of known routes that will not be part of the scope of the risk assessment. For example, in many waterborne pathogen risk assessments, the ingestion route of exposure is investigated and other routes of exposure (e.g., inhalation) are not included. In cases where a pathogen is not known to be infectious through certain routes, such as the dermal route, discussion, including rationale, for not including the dermal exposure route should be included in the risk assessment documentation, particularly for risk assessments conducted to meet specific statutory requirements where reasons for excluding a route must be justified. Additional discussion of how the choice of included routes impacts the uncertainty, qualitative or quantitative, should be included in the risk characterization section.

For an infection to occur, the host's target organ must come in contact with a sufficient number of microorganisms; the microorganism must possess specific virulence factors; these virulence factors must be expressed; and the defenses of the host and/or target organ systems (e.g., digestive system, lung) must be overcome. With some microorganisms (e.g., *Giardia*, *Cryptosporidium*), the interaction with the particular organ is so specific that infections are almost always confined to that one organ site; with others (e.g., *Salmonella*, enteroviruses) the pathogen has the potential to infect more than one target organ. When attempting to establish a health risk due to exposure to pathogens through contact with food and drinking water, one must consider that the human GI tract is a complex organ system with a variety of specific host defense mechanisms. It is only when

[12] The term "vector" can mean "anything which transmits parasites" (for use in this MRA Tools document, a vector can transmit bacteria, viruses, or parasites) (www.swintons.net/jonathan/Academic/glossary.html) or can refer to intermediate hosts that are required for life cycle completion. Environmental reservoirs include free-living amoeba that can harbor bacteria intracellularly allowing the bacteria to survive in harsher environments than they could normally survive (NRC, 2004).

[13] Route of exposure refers to how the pathogen comes in contact with the vulnerable host receptor cells that support infection (e.g., inhalation, dermal contact, oral), whereas source of exposure refers to the physical matrix that carries the pathogen (e.g., air, water, food, soil).

the pathogen has particular virulence factors for sites in the GI tract, and the specific host defense mechanisms in the GI tract are breached, does infection occur. Infection without symptoms and the duration of infection are important attributes of the infection process because they contribute to the potential for secondary transmission via the shedding of pathogens into the environment.

An MRA can also include exposures that have not previously been commonly considered. For example, Feazel et al., (2009) analyzed ribosomal ribonucleic acid gene sequences from 45 showerhead sites around the United States. The majority of showerhead microbiota were comprised of genus- or species-level groups that are commonly found in water and soil. The showerhead environment strongly enriches for microbes that are known to form biofilms in water systems, including *Mycobacterium* spp., *Sphingomonas* spp., *Methylobacterium* spp., and other pathogens. The detection of significant loads of *M. avium* in showerhead biofilms identifies a potential personal health concern. *Legionella*, which can cause Legionnaires disease and Pontiac fever, has been linked epidemiologically to hot water systems, cooling towers, evaporative condensers, humidifiers, whirlpool spas, respiratory devices, and decorative fountains. It has also been isolated from those sources as well as water taps, hot tubs, showers, creeks, ponds, and soils from the banks of water bodies (APHA, 2004). There may also be other opportunistic pathogens that are of interest to risk managers and assessors.

Secondary Transmission

The potential for secondary transmission will also contribute to human exposure. Secondary transmission refers to infection spreading from one infected person to another person. Secondary cases (often represented by a secondary attack rate) generally refer to cases or an attack rate that occurs among contacts, within the incubation period of the pathogen, and following exposure to a primary case. In some cases, direct person-to-person transmission cannot be separate from contamination of the immediate environment and subsequent transmission to another person (e.g., toddlers sharing toys versus direct physical contact during play). In most cases, it is appropriate that the definition of secondary transmission include infections that result from propagation of the specific exposure of interest, but not encompass distant transmissions (separated by time and/or space) that may be more appropriately considered to result as a function of person-to-environment-to-person transmission. Temporal and spatial limitations should be specifically noted in the definition of secondary transmission for a given pathogen. Full discussion of the range of scenarios that qualify as secondary transmission should be included where appropriate.

The above definition of secondary transmission is limited to avoid overlap with pathogen occurrence in the environment (person-environment-person), although people are, of course, part of the environment. However, the potential for re-introduction of the pathogen into the exposure media can also be within the definition of secondary transmission. Dynamic MRA models can characterize secondary cases that occur among contacts following exposure to a primary case, whereas static MRA models usually consider secondary transmission to be negligible or include it as a non-fluctuating multiplicative factor (e.g., secondary cases equal primary cases multiplied by 0.1; assuming a 10% secondary transmission rate). The problem formulation documentation should indicate if and how secondary transmission is included in the assessment. If it is not included, justification for this decision should be provided.

Taxonomy and Strain Variation

Taxonomy and strain variation have a potentially large effect on risk assessment. The difference in dose-response range between isolates (and strains) can be orders of magnitude (see Table 5). Some strains may not be infective for humans. In addition, the ratio of different strains in the environment can fluctuate. These factors make characterization of pathogen occurrence difficult (Messner et al., 2001; Teunis et al., 2002). The extent to which strain variation is accounted for in the risk assessment should be documented during problem formulation.

Table 5. Virulence of Three *Cryptosporidium parvum* Isolates in Healthy Adult Humans (Source: Okhuysen et al., 1999)

	Isolate		
Isolate Characteristic	**Iowa**	**UCP**	**TAMU**
Infectious dose for 50% of population (ID_{50})	87 oocysts	1042 oocysts	9 oocysts
Attack rate	86%	52%	59%
Duration of symptoms	64.2 hours	81.6 hours	94.5 hours

2.4.2. Initial Host Characterization

Host characterization involves an evaluation of the intrinsic and acquired traits that modify the risk of infection or illness in a potentially exposed human population. It is also possible that host factors may be important in determining the severity or outcome of an infection. For example, high-risk groups may develop severe symptomatic illness, whereas, low-risk groups may develop asymptomatic infections or mild illness.

The following populations are typically considered more susceptible[14] than the general population: pregnant women; neonates and children; people over 65 years of age; individuals residing in nursing homes or related care facilities; and cancer, organ transplant, and acquired immune deficiency syndrome (AIDS) patients (Haas et al., 1999). The Report to Congress, *EPA Studies on Sensitive Subpopulations and Drinking Water Contaminants* (U.S. EPA, 2000d), summarizes EPA's approach to identifying and characterizing susceptible subpopulations that may be at greater risk from exposure to drinking water contaminants than the general population.

Host characteristics have the potential to influence both the exposure and the health effects components of the risk assessment. These factors are often used to define potential subpopulations of interest for a risk assessment because they can influence the assessment with respect to

[14] Sensitive subgroups are "identifiable subsets of the general population that, due to differential exposure or susceptibility, are at greater risk than the general population to the toxic effects of a specific air pollutant (e.g., depending on the pollutant and the exposure circumstances, these may be groups such as subsistence fishers, infants, asthmatics, or the elderly)" (U.S. EPA, 2007a). Susceptible subgroups "may refer to life stages, for example, children or the elderly, or to other segments of the population, for example, asthmatics or the immune-compromised, but are likely to be somewhat chemical-specific and may not be consistently defined in all cases" (U.S. EPA, 2007a). Note that in the above definitions "susceptible" refers to host characteristics and "sensitive" refers to host characteristics and exposure patterns. Another similar usage of terms is "susceptible" to refer to intrinsic factors and behavioral and situational factors as extrinsic, with both intrinsic and extrinsic being grouped under "vulnerable."

susceptibility to infection and severity of illness. The extent to which these factors are considered in the risk assessment should be described in the problem formulation documentation. Factors related to host characterization that should be considered during problem formulation are briefly discussed below (also see Table 1):

- Part of host characterization in problem formulation is defining demographics of the populations of concern. The size of the exposed population refers to the number of people who come in contact with the media of concern. The demographics and behavior of the exposed population can conceptually include many possible subgroups. Defining the subpopulations that will be considered is a key component of problem formulation. Subgroup differentiation is not necessary unless there is evidence for relevant differences between the subgroups.[15] There should be scientific rationale presented for dividing subgroups as well as data that directly pertain to that subgroup or can be adjusted to address that subgroup. For example, it is unlikely that differentiating between 24 and 25 year-olds would provide any additional useful information for risk managers.
- The young and the elderly generally have less resistance to infections[16]. Children, especially malnourished children, may be more likely to exhibit severe effects of acute GI illness (AGI) after exposure to some pathogens (e.g., pathogenic *E. coli*, some enteric viruses). However, some pathogens (e.g., Hepatitis A, poliovirus) may cause less clinical illness in children than in adults (Gerba et al., 1996a). Age can also contribute to different exposure patterns due to behavior. For example, children may have higher levels of incidental ingestion of water during swimming than adults (Dufour et al., 2006). Because drinking water consumption increases with age, the elderly consume more drinking water than adults or children (Roseberry and Burmaster, 1992).
- Populations that are considered immunocompromised or immunosuppressed due to recent or concurrent illness or medical treatment may be defined as subpopulations that the risk assessment will address (Effler et al., 2001). However, all definitions of subpopulations included in the risk assessment should include the criteria used to classify individuals as immunocompromised and may need to be limited to specific identifiable types of immune defects. Extreme physical or emotional stress can lower immune competency. The host GI environment can vary in ways that affect pathogens and innate immunity also plays a role in infection dynamics.
- Previous exposure may confer limited and/or short-term protective immunity for some pathogens (Frost et al., 2005). The converse of this may also be true; that is, when individuals or populations that have not previously been exposed to particular pathogens, infection and illness rates can be higher than would otherwise be anticipated. "Traveler's diarrhea" is a well-known observed phenomenon that exemplifies this type of situation.
- Concurrent illness/medical treatment (physical and mental stressors may increase susceptibility).

[15] Note that risk assessments being performed as part of a statutory requirement may already have mandated subgroups.

[16] Although children are referred to in conjunction with subpopulations in this document, U.S. EPA acknowledges that childhood represents a life-stage rather than a subpopulation, the distinction being that a subpopulation refers to a portion of the population, whereas a life-stage is inclusive of the entire population (http://www.U.S. EPA.gov/teach/index.html).

- Genetic background can also affect immune status, but may play a larger role in mechanism of infection and disease progress.
- Pregnancy may cause women to be more susceptible to a pathogen. For example, Hepatitis E, which causes a self-limiting disease in most infected persons, can cause up to 20% mortality in women in the third trimester of pregnancy (Jameel, 1999).
- Malnourished individuals tend to have weaker immune defenses than well-nourished individuals.
- Social and behavioral traits primarily affect exposure patterns. For example, a relatively small proportion of the population is responsible for consuming the majority of raw and partially cooked shellfish (FDA, 2005). As mentioned above, age may also be related to behaviors that affect pathogen exposure patterns.

Data for the above elements can be arranged into groups by stratification or multivariate analysis. Alternatively, host characteristics can be considered by conducting a separate risk assessment for each characteristic that is believed to have some importance. For example, in addition to a risk assessment for the overall population, a separate risk assessment may be performed for each subpopulation of interest (e.g., young children, the elderly, pregnant women, immunocompromised persons) provided that sufficient data are available for valid statistical interpretation. EPA's Risk Assessment Forum has developed *Guidance on Selecting Age Groups for Monitoring and Assessing Childhood Exposures to Environmental Contaminants* (U.S. EPA, 2005b), which recommends subgroups that address anatomy/physiological development in the following age groupings: birth to <1 month, 1 to <3 months, 3 to <6 months, 6 to <12 months, 1 to <3 years, 3 to <8 (female) or <9 (male) years, and 8 or 9 years to <16 (female) or <18 (male) years.

In recognition that children have a special vulnerability to many toxic substances, the EPA Administrator's October 1995 *Policy on Evaluating Health Risks to Children* directs the Agency to explicitly and consistently take into account environmental health risks to infants and children in all risk assessments, risk characterizations, and public health standards set for the United States. In April 1997, President Clinton signed Executive Order (EO) 13045 *Protection of Children from Environmental Health Risks and Safety Risks*, which assigned a high priority to addressing risks to children (EO, 1997). In May 1997, EPA established the Office of Children's Health Protection to ensure the implementation of the President's EO. EPA has increased efforts to ensure its guidance and regulations take into account risks to children. In 2002, EPA published an interim report on child-specific exposure factors (U.S. EPA, 2002d).

2.4.2.1. Environmental Justice

EO 1289, *Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations* (February 1994), ordered Federal agencies, including EPA, to "…make achieving environmental justice part of its mission by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of its programs, policies, and activities on minority populations and low-income populations…" (EO, 1994). EPA responded to EO 12898 with *The EPA's Environmental Justice Strategy* (U.S. EPA, 1995b). In 2001, in a memo from the EPA Administrator environmental justice was defined as follows: "The Agency defines environmental justice to mean the fair treatment of people of all races, cultures,

and incomes with respect to the development, implementation, and enforcement of environmental laws and policies, and their meaningful involvement in the decision-making processes of the government" (emphasis in original).[17] EPA further defined meaningful involvement in EPA's *Public Involvement Policy* (U.S. EPA, 2003g) as follows:

> "Meaningful involvement"…means that: (1) potentially affected community residents have an appropriate opportunity to participate in decisions about a proposed activity that will affect their environment and/or health; (2) the public's contribution can influence the regulatory agency's decision; (3) the concerns of all participants involved will be considered in the decision-making process; and (4) the decision-makers seek out and facilitate the involvement of those potentially affected.

Risk assessment documentation should provide clear descriptions of subpopulations and other parameters that may help EPA evaluate whether there are potential environmental disparities that could cause an environmental justice concern.

2.5. Linkage between Problem Formulation and Other MRA Components

The planning and scoping and problem formulation process develops the scope of the risk assessment, taking into account management needs, Agency risk assessment policies, risk assessment tool availability, data constraints, and overall Agency resources. The output from this process—documentation of the problem formulation development—provides the linkage to rest of the MRA process.

A problem formulation lays out how everything fits together in the risk assessment. Well-formulated problem formulation effectively communicates the scope, purpose, and methods that will be used in the MRA. It provides clarity about the stressor and provides a conceptual model for assessing risk. Finally it provides an analysis plan describing how exposure will be evaluated, how the hazard will be evaluated, and how these will be integrated to develop the risk assessment.

[17] http://earth1.EPA.gov/oswer/ej/html-doc/ejmemo.htm

3. Exposure Assessment

The characterization of exposure consists of the technical evaluation of data related to the potential exposure to microbial contaminants in water media. The problem formulation phase of the MRA precedes the exposure assessment and may partially address many of the issues to be evaluated in the exposure assessment. However, the exposure assessment is more detailed and generally more quantitative than the problem formulation phase.

In human health risk assessment, *exposure* is defined as human contact with a biological, physical, or chemical agent—usually through ingestion, inhalation, or dermal contact. Risk assessment can be performed for specific target populations or an individual target organism (a human with a defined exposure pattern). Exposure assessment involves the determination or estimation (qualitative or quantitative) of the magnitude, frequency, duration, and route(s) of exposure (U.S. EPA, 1997a). A primary purpose of exposure estimation is to support dose estimation (U.S. EPA, 1992). *Dose* is the amount of a pathogen that enters or interacts with a host (ILSI, 2000).[18]

For nearly all MRA contexts, dose refers to potential dose (i.e., the number of pathogens ingested in a specified period) because the actual numbers of pathogens that an individual is exposed to is almost always unknown. In fact, most MRAs are performed without direct estimates of pathogen dose. Doses are typically calculated as a function of pathogen density in the exposure medium (e.g., drinking water, reclaimed water, biosolids) and the volume of that medium that is ingested or inhaled. An important reason for calculating pathogen doses is that doing so allows the data from one exposed population (e.g., the volunteers in a virulence study) to be applied to risk assessments for other exposed groups, such as the general population.

Characterization of exposure involves an evaluation of the interaction between the pathogen, the environment, and the human population (i.e., the classic epidemiological triad; Figure 9). The infectious disease hazard characterization, occurrence, and exposure assessment sections are brought together to develop an Exposure Profile that quantitatively or qualitatively summarizes the magnitude, frequency, and pattern of human exposure for the scenario(s) under investigation.

Exposure is not limited to a pathogen-specific context. It can also be defined in terms of water quality indicators such as the presence of coliforms, pathogen surrogates, or types of water sources (e.g., ground water, impoundments, rivers) coupled with estimated efficiencies of treatment

[18] U.S. EPA (U.S. EPA, 1997a, 2003e, 2004b, 2005a) has defined "*dose* as the amount of a substance available for interactions with metabolic processes or biologically significant receptors after crossing the outer boundary of an organism. The *potential dose* is the amount ingested, inhaled, or applied to the skin. The *applied dose* is the amount presented to an absorption barrier and available for absorption (although not necessarily having yet crossed the outer boundary of the organism). The *absorbed dose* is the amount crossing a specific absorption barrier (e.g., the exchange boundaries of the skin, lung, and digestive tract) through uptake processes. *Internal dose* is a more general term denoting the amount absorbed without respect to specific absorption barriers or exchange boundaries. The amount of the chemical available for interaction by any particular organ or cell is termed the *delivered* or *biologically effective dose* for that organ or cell." Note that these sub-definitions of dose are not used to describe pathogen interactions with hosts. The term "minimum infectious dose" was intended to indicate the lowest dose that would cause infection in an individual and assumed that there was a threshold dose. This term is generally considered to be obsolete because as little as one microorganism is believed to be capable of causing infection in a susceptible individual. However, it should be reiterated that infection does not imply symptomatic illness (U.S. EPA, 2007a).

technologies (e.g., filtered and disinfected water versus disinfected water only). In all cases where indicators are used in risk assessment, it is important to document fully the basis for their use (i.e., the extent to which they are correlated with the pathogen or health effect of concern), and to specify clearly the conditions under which the correlation is expected to be valid. QMRA for indicators and surrogates is beyond the scope of this document. For more information on using this type of information, please refer to EPA's Technical Support Materials (U.S. EPA, 2014).

Among elements of MRAs, the exposure assessment is the one with the greatest flexibility in its formulation and often, the greatest data needs. Although some MRAs may be created to provide large-scale or general estimates of microbial risk (e.g., nationwide estimates of GI illness due to consumption of drinking water), most QMRAs estimate risk for specific venues and pathogens. As noted above, sources, transport processes, and epidemic and episodic occurrences of pathogens are highly variable both at a particular site and between sites. In exposure assessment, an exposure profile is developed, inclusive of all the processes, variabilities, and uncertainties for the venue of interest. Thus, the exposure assessment must assemble models and data that address the questions such as following:

- What are the sources of the pathogens?
    - What is the temporal variation of the occurrence and abundance of the pathogen of interest in the source?
    - How certain are measurements of the pathogen density in the source?
- What is the pathway from source to receptor over which the pathogen must survive and be transported?
    - What are the transport or treatment processes that determine the microorganism density at the point of ingestion?
    - How do those processes vary with time?
    - How persistent is the microorganism in the water matrix and for the physical conditions at the venue of interest?
- What is the route of exposure (inhalation, oral, dermal)?
    - If oral, what volumes of water will be ingested and where will ingestion occur?
        - Do ingestion rates vary among potentially affected population groups?
        - How do sites where ingestion occurs differ?
- What are the magnitude, duration, and frequency of the exposures of interest?

The answers to these types of questions will be quite different for different hazards. Sites and models describing exposure may include highly detailed flow models, as included in an assessment of the benefits of additional wastewater treatment (Soller et al., 2003). They can be simple, but effective, as in a study of cryptosporidiosis arising from drinking water consumption by Pouillot et al. (2004). In all MRAs, exposure assessments should be sufficiently detailed to meet the MRA objectives and be consistent with the conceptual models generated in problem formulation. Exposure assessment data needs are determined by the MRA objectives and by the formulation of the exposure model. As described in Section 5.3, there are advantages to characterizing some parameters of the exposure model as statistical distributions. Selection and parameterization of those distributions require data. The more data that are generated or gathered, the more certain the distribution choice and parameters will be.

3.1. Occurrence

Occurrence refers to the conditions that lead to the presence of a pathogen or to the distribution pattern of a pathogen in the environment and the media of concern. The EPA-ILSI Framework (ILSI, 2000) identifies the following elements to consider when characterizing pathogen occurrence during exposure assessment:

- temporal distribution/frequency;
- density in environmental media;
- spatial distribution (clumping, aggregation, particles, clustering);
- niche (ecology, non-human reservoirs);
- survival, persistence, amplification;
- seasonality;
- meteorological and climatic events; and
- presence and effectiveness of treatment or control processes.

Many of these factors are interrelated and as such cannot be discussed independently. There are three basic questions that should be answered to describe pathogen occurrence in a water body—when (including duration), where, and how much (level)? When information on a particular pathogen species of interest is lacking, it may be necessary to use occurrence data for surrogate or index species. The limitations and uncertainty associated with those data and their use should be evaluated and discussed.

3.1.1. When Do Pathogens Occur in the Water Body?

Temporal distribution/frequency describes when pathogens occur. Fluctuations in microbial densities can occur on almost any time scale (Boehm, 2007; Boehm et al., 2002) and over a wide range of spatial scales. In many cases, pathogen occurrence can vary on wide-ranging time scales, including hourly, daily, weekly, monthly, seasonally, or yearly fluctuations. Spatial variations often relate to the position of a pathogen source relative to the location of receivers or to mixing processes at a site of interest. Representative drivers for these fluctuations are listed along with their associated length and time scales in Table 6.

Meteorological and climatic events such as storms, changes in wind direction and shifts in currents may cause changes in pathogen occurrence. Seasonality is a factor that affects the temporal frequency of many waterborne pathogens, such as *Cryptosporidium*. Seasonality also influences microbial density in environmental waters—either because of dependence of persistence on temperature or because of differences in loading between wet and dry seasons. Moreover, seasonal animal-related events, such as calving or bearing young seem to be associated with some zoonotic pathogens. Occurrence data that are linked to temporal events such as seasons may be useful for predicting how pathogen levels may respond to future events. If wastewater treatment or control processes are less efficient, as may occur during storm events (resulting in combined sewer overflows and sanitary sewer overflows), there may be associated temporal fluctuations in pathogen levels. Urban and agricultural runoff can also influence pathogen occurrence in surface waters.

Table 6. Sources and Scales of Variability in Pathogen Occurrence

Variability Type	Variability Driver	Length Scale	Time Scale
Temporal	Seasonal changes	Regional	Months
	Human or animal epidemics	Watershed or sub-watershed	Weeks to months
	Hydrometerological events	Watershed	Days
	Treatment plant operational events	Watershed for WWTPs, distribution system for WTPs	Hours to days
	Animal life-cycles	Sub-watershed	Weeks to months
	Tidal process	—	Hours
	Solar cycle	—	Hours
	Wave action or bottom scour	Meters	Minutes
Spatial	Alignment of source with water of interest	Kilometers	—
	Hydraulics and mixing	Meters to kilometers	Minutes to days
	Distribution of sources	Watershed	—

Disease epidemics in a community also can affect pathogen occurrence, particularly in wastewater and wastewater-impacted water bodies. Consideration of fluctuations in endemic levels of disease in a community has the potential to strongly affect the interpretation of a risk assessment; therefore, the characterization of exposure should be explicit in terms of whether endemic or epidemic conditions, average or peak flow events, or specific events are to be evaluated, and a justification for that decision should be provided.

The need to measure pathogens at low densities and the episodic nature of pathogen loading are limitations of direct pathogen monitoring or the exclusive use of pathogen data in MRA exposure assessments. For some of the pathogens of interest in waterborne exposures, very low numbers of organisms can result in a high probability of infection. Despite advances in microbial monitoring and particularly the use of molecular methods, detection of pathogens at levels of interest may not be possible for some of the pathogens of concern or in some of the water matrices of interest (NRC, 2004). Difficulties in direct measurement of pathogen densities are particularly acute in drinking water treatment processes whose intended effect is reduction of pathogens to very low densities. These difficulties suggest the following two strategies for estimating exposure to pathogens:

- using distributions of pathogen densities in circumstances when relatively high densities are known to be present along with modeling to determine pathogen densities at a point of interest, or
- basing exposure on a surrogate measure of microbiological quality.

Fecal indicator organisms may be used in in development of criteria and assessment of the safety of waters used for recreation. Development of criteria values for fecal indicator organisms is beyond the scope of this document. The following information is presented to show how fecal indicators are used as surrogates for the occurrence of pathogens. Epidemiological studies have been used to establish relationships between fecal indicator organism density and human health risk for sites with known fecal pollution sources (Parkhurst et al., 2007). Here, it is accepted that

waters from specific fecal pollution sources pose a characteristic risk to exposed persons. Given this assumption, the fecal indicator organism density indicates the extent to which fecal pollution from the specific source is present and is related to the risk of illness or infection at the site. Health effects relationships based on indicator densities may be used directly in MRA illness estimation when the fecal pollution source at the site of interest is the same as that for which the health effects relationship was developed and with the understanding that the health effects relationship depends on the indicator density, not the ingested indicator dose. Alternatively, the indicator density and expected adverse health effects may be used to estimate pathogen densities in a known fecal pollution source, as in a "reverse QMRA" process (U.S. EPA, 2014; Soller et al., 2014; Soller et al., 2010a).

The second potential role for indicator organisms in QMRA is in characterization of process efficiency, as in the use of indicator organisms in assessment of drinking water treatment risks. For example, total coliforms do not pose a health hazard and are not related to a specific fecal pollution source; however, they are relatively abundant at stages of the drinking water treatment process where pathogens, if present at all, may be present at non-detectable densities. Total coliform removal in unit processes may be characterized by a distribution or a range of values. Comparison of total coliform removal for a specific unit process at a given time to the range of removals for the process provides an indication of how well the process is operating. Because removal rates differ among microorganism, the removal of total coliforms cannot be assumed similar to that of pathogens. Rather, they may indicate where, within a removal range for a given pathogen, the unit process performance may lie or whether water treatment failure has occurred and whether a revised performance range should be used.

3.1.2. Where Do Pathogens Occur in the Water Body?

Spatial distribution of pathogenic microorganisms can differ depending on the microorganism and on the properties of the water matrix. If pathogen occurrence fluctuates over time, then the degree of clumping, aggregation, and clustering may also change as water parameters change. Unlike chemicals, pathogens are particulates and may stick to each other or to sediment and other particles (Gerba et al., 1991). The size and nature of particles will influence suspension and settling in different hydraulic conditions. Therefore, particles that carry pathogens may be distributed within a water body in an uneven (heterogeneous) manner. For waterborne pathogens, niche is relevant for "free living" species. Pathogens may thrive in open water, sediments, or other ecologically defined spaces. Non-human reservoirs may also be an important part of the pathogen ecology of a particular pathogen. If there is appreciable survival, persistence, or amplification in non-human species, then those sources of contamination of water may need to be considered during the occurrence assessment. This includes animals (wildlife and domestic) as well as other microorganisms. Survival, persistence, and amplification can differ in different microclimates within a water body, and should be considered factors that influence where pathogens occur. Pathogens in sediments may be resuspended in the water column due to changes in flows associated with precipitation, runoff, tides, and currents.

3.1.3. What is the Level of Pathogens in the Water Body?

Characterizing pathogen occurrence relies on measuring the density in the environment and correlating density with spatial and temporal patterns in the environment, such as niches, seasons, weather events, and human-related activities. There are several difficulties that are commonly encountered when measuring pathogen levels in water samples. Because microorganisms tend to clump and aggregate (heterogeneous distribution), replicate samples can yield measurements that differ substantially (even by orders of magnitude) (see Section 3.1.4 below).

The ability for a pathogen to survive and also remain infectious in a water body is dependent on both pathogen characteristics and environmental factors. Pathogen-specific characteristics include but are not limited to, genetic strain variations, the growth conditions the pathogen experienced before entering the water body, duration in the environment, protective states, and VBNC states. Environmental factors include but are not limited to, temperature, pH, turbidity, nutrient levels, osmotic conditions, ultraviolet light exposure, predation, and interactions with other living organisms. For example, and as noted previously, amoebae may be reservoirs that may contribute to the survival, persistence, and/or amplification of environmental pathogens. Fate and transport modeling can provide plausible scenarios and estimates of how microbial densities can change over time as they move through the aquatic environment.

Pathogen occurrence patterns will also be affected (but not necessarily in a similar manner) by the presence of control strategies and treatment processes (either wastewater or drinking water treatment depending on the context). Mitigation strategies may involve improving existing control processes or adding new control measures, which can be modeled in the risk assessment. Discussion of the sources of microbes may be helpful in characterizing occurrence patterns. Some commonly considered sources include wastewater treatment plant effluent, some industrial effluents, leaking septic tanks, urban runoff, agricultural runoff, animals (e.g., livestock, domestic, wildlife), and environmental niches (e.g., sediments, aquatic plant life). Densities of pathogens vary in untreated sewage based on the level of shedding in the contributing human population and densities in treated sewage vary based on levels before treatment and efficacy and type of treatment processes. Differences in contributing populations can result in orders of magnitude differences in microbial levels in sewage (Gerba et al., 2008).

Pathogens occur at different densities in water and wastewater treatment processes, whose primary function is reduction of the densities of pathogens and other contaminants. Microbial risk assessment provides a framework in which the densities of pathogens, from source to exposure, may be estimated. This ability of MRA is important, given the variability in occurrence and density of pathogens in different source waters, the variability in performance of unit operations, the serial nature of water and wastewater treatment, and the dependence of performance of unit operations on prior operations.

The technical literature contains numerous studies characterizing the range of unit process operation removals for different pathogens and different operations (e.g., Betancourt and Rose, 2004, for removal of *Cryptosporidium* and *Giardia* in drinking water treatment processes). Other studies report the overall removal of pathogens or their surrogates for entire water treatment processes (e.g., Castro-Hermida, 2008, for removal of *Giardia* and *Cryptosporidium* from multiple

drinking water treatment plants; or Kistemann et al., 2008, for removal of multiple pathogens or their surrogates in drinking water treatment plants). These studies demonstrate both the potential for generating general ranges of removals that might be expected in different unit operations and the variability of the efficiencies between plants and for different water matrices.

Numerous QMRAs have been conducted of drinking water treatment processes, several of which are used as illustrations in other sections of this report (Olivieri et al., 1999; Westrell et al., 2003; Pouillot et al., 2004; Petterson et al., 2007). An additional study (Smeets et al., 2008) illustrates the use of ranges or distributions of removals for unit process operations within a model capable of predicting risk for a particular treatment train. In that study, yearly *Campylobacter* illness risks from consumption of drinking water produced at a plant employing filtration and ozonation were estimated. Several parameters in the model constructed for this process are variable, including raw water *Campylobacter* density, removal efficiency for filtration, and removal efficiency for ozonation. Smeets et al. (2008) evaluated multiple approaches for developing distributions to characterize the removal processes. Removal data for filtration and ozonation were based either on paired data (influent and effluent densities of the unit operations) taken at the same day, or data paired based on their rank. Additionally, both parametric and non-parametric distributions were used for removal efficiencies, with the distributions based on the removal efficiency data. The authors evaluated Weibull, gamma, and log-normal distributions and determined that the raw water *Campylobacter* density data, filtration removal efficiency data, and ozonation removal efficiency data were all fit best by log-normal distributions. The authors concluded the following from their study:

- the rank method provided much better agreement between predicted and observed *Campylobacter* densities for the unit operations;
- use of parametric representations of the removal efficiencies of the unit operations is superior to non-parametric representations in this instance because parametric models allow for the occurrence of rare events, although use of non-parametric methods reduces the effect of distribution choice on risk estimation; and
- use of QMRA in assessing drinking water risk complements pathogen monitoring, especially given the low densities of pathogens occurring in parts of the water treatment process.

The outcome of the "occurrence" section of the process is an evaluation of all relevant factors pertaining to the occurrence and distribution of the pathogen. Several tools and databases for evaluation of occurrence, which may be useful for MRA exposure scenarios, are summarized in Table 7.

Table 7. Tools and Databases for Evaluation of Occurrence

Tools	Reference(s)
EPA STORET and WQX	http://www.epa.gov/storet/about.html The STORET Data Warehouse is EPA's repository of the water quality monitoring data collected by water resource management groups across the country. Data can then be re-used for analysis. WQX is the framework by which organizations submit data to the Warehouse.

	Table 7. Tools and Databases for Evaluation of Occurrence (continued)
EPA Basins Program	http://www.EPA.gov/OST/BASINS/ A multi-purpose environmental analysis system that integrates a geographical information system, national watershed data, and state-of-the-art environmental assessment and modeling tools into one convenient package. Download at http://www.EPA.gov/waterscience/basins/basinsv3.htm. This release includes additional links to water quality models as well as a new data user interface tool with access to national data layers.
EPA EMPACT (Environmental Monitoring for Public Access and Community Tracking) Study	General information is available at: http://www.EPA.gov/nerl/news/forum2003/water/brenner_poster.pdf and http://www.EPA.gov/ORD/NRMRL/pubs/625r02017/beaches_html/chapter1.html. Location-specific EMPACT studies are available on the Internet.
EPA Information Collection Rule (ICR) and Supplemental Surveys (provides summarized data on *Cryptosporidium* only)	Overview with links for further information is available at: http://www.EPA.gov/enviro/html/icr/ and http://www.EPA.gov/safewater/icr.html.
EPA's Unregulated Contaminant Monitoring Program	EPA uses the Unregulated Contaminant Monitoring program to collect data for contaminants suspected to be present in drinking water, but that do not have health-based standards set under the Safe Drinking Water Act (SDWA). http://water.epa.gov/lawsregs/rulesregs/sdwa/ucmr/
Safe Drinking Water Information System (SDWIS)	SDWIS—Federal (SDWIS/FED) version is U.S. EPA's national regulatory compliance database for the drinking water program. It includes information on the nation's 170,000 public water systems and violations of drinking water regulations. • Access Drinking Water Information Online (through summary pivot tables, Envirofacts, or direct connection to the mainframe) http://www.EPA.gov/safewater/data/getdata.html. • SDWIS/FED Website (information for users who work with the database) http://www.EPA.gov/safewater/sdwisfed/sdwis.htm.
Statistical method: Markov Chain Monte Carlo (MCMC) simulation for modeling environmental pathogen densities in natural waters	Crainiceanu et al. (2003) "Modeling the United States national distribution of waterborne pathogen concentrations with application to *Cryptosporidium parvum.*"
AWWARF report on effects of meteorological events (*Cryptosporidium* only)	http://awwarf.org/research/topicsandprojects/execSum/488.aspx. "Rainfall events and other watershed perturbations, especially those during the spring runoff, pose the greatest risk for causing waterborne cryptosporidiosis."
USDA agricultural runoff models (hydromodels)	http://wmc.ar.nrcs.usda.gov/technical/WQ/modeldesc.html Information and links on 15 Natural Resources Conservation Service models related to water and agriculture.
USDA/ARS HYDRUS models	http://www.ars.usda.gov/Services/docs.htm?docid=8910 Simulates water flow and solute transport in a two-dimensional variably-saturated medium (graphical user interface).
Surfrider	The Blue Water Task Force is the Surfrider Foundation's volunteer-run water testing program. http://www.surfrider.org/blue-water-task-force

3.1.4. Interpretation of Analytical Methods

It is important to document measurement techniques and their capabilities and limitations carefully so that scientifically defensible decisions can be made about integrating or not integrating results from different studies. For example, there may be differences in the relative ratio of infective, viable, and nonviable *Cryptosporidium* oocysts in the samples taken from source water versus laboratory generated sources of oocysts used in human trails. Young, freshly harvested oocysts are typically used for the human trials and are most often used for disinfection studies, yet when conducting environmental sampling, the age of the oocysts in any given sample can vary widely, as could the distribution of viable and infective oocysts. This difference leads to part of the uncertainty in the interpretation of the data, and thus assumptions regarding how laboratory generated data are applied to environmental scenarios need to be clearly stated in exposure assessment.

For data sets that involve laboratory methods issues related to sensitivity, specificity, limit of detection, sampling method, and sample size should be examined. Their effects on the risk assessment assumptions should be discussed. Culture-based approaches rely on growing the organism in question, isolating it in pure culture, and characterizing the morphological, biochemical, physiological and other traits. However, not all microorganisms are culturable, such as VBNC, and some parasites and viruses.

Molecular methods can allow serotyping and determination of virulence traits. The risk assessment documentation should discuss any relevant information related to analytical methods, including error bars if possible. For example, Dupont et al. (1995) included standard deviations for the actual numbers of *C. parvum* oocysts given experimentally to subjects versus what the intended dose was. For an intended dose of 30 oocysts, the actual dose was 34 ± 3, for 100 oocysts the actual dose was 108 ± 22, and for 300 oocysts, the actual dose was 313 ± 24. When new methods are introduced, such as cell culture techniques to enumerate the number of infectious oocysts in water samples, they may allow refinement of the dose-response relationship (Slifko et al., 1997, 1999, 2002). This refinement can then be used to improve future exposure assessments.

EPA's LT2 has a discussion of EPA Methods 1622/23 performance and how level of performance influences the occurrence data for the Information Collection Rule (ICR) (U.S. EPA, 2006a). To estimate how these errors would affect the assignment of systems to categories (bins) with different levels of occurrence, EPA constructed a Monte Carlo model that dealt with the error components in the following manner:

- Finite volume assayed—the model defines the number of oocysts present in a 10 L volume as a Poisson random variable, whose mean is the product of measurement recovery, volume assayed, and density at the time of sampling.
- Finite number of samples—true density varies over time as a random variable. Density is modeled to vary in such a way that its natural logarithm is normally distributed with standard deviation 1.762. This value was selected based on Bayesian analysis of survey data and on expert opinion that at any given site the *Cryptosporidium* density would vary within a three order of magnitude density range 95% of the time (i.e., 2.5% of the time the density would be less than X, and 2.5% of the time the density would exceed 1000X).

- Variable recovery—based on laboratory performance in ICR Supplemental Survey, EPA assumed for the model an average recovery among all laboratories of 40% with a relative standard deviation of 50%. Recovery is modeled as a Beta random variable with parameters $(\alpha, \beta) = (2, 3)$. Mean recovery is therefore $\alpha/(\alpha + \beta) = 2/(2 + 3) = 0.4$. The standard deviation of recovery is 0.2, which is half the mean recovery.

3.2. Exposure Analysis

An exposure scenario summary can be a short narrative description of how an individual is exposed to a hazard. A more formal exposure scenario provides additional detail about the range of exposures that are considered in the risk assessment. The EPA *Exposure Factors Handbook* (U.S. EPA, 1997a, 2011) is the Agency-wide resource for building exposure scenarios for chemical hazards. It can also be consulted for data that may apply to infectious disease hazard exposures. The interagency *Microbial Risk Assessment Guideline* has considerable detail on exposure models (U.S. EPA/USDA, 2012). Elements that should be evaluated for inclusion in exposure assessment and are used to define the exposure scenario scope are presented below (adapted from ILSI, 2000):

- identification of media;
- routes of exposure (including secondary transmission);
- units of exposure (period of relevancy to characterize dose, includes magnitude, duration and frequency);
- temporal nature of exposure (whether single or multiple exposures);
- spatial nature of exposure; and
- behavior of exposed population.

A summary of parameters used in an example EPA exposure analysis is presented in Text Box 4.

Text Box 4. Exposure Analysis for the Long Term 2 (LT2) Enhanced Surface Water Treatment Rule (Source: U.S. EPA, 2006a).

Media – the media considered was surface water used as a source for drinking water from filtered systems and unfiltered systems.

Units of Exposure (magnitude, duration, frequency) – # of oocysts per ingestion volume (magnitude), per day (duration), for an annual frequency of 21 days (frequency)

Ingestion Volume – individual consumption mean = 1.071 L per day; (LT2 page N-16)

Frequency of Exposure – annual days of exposure = 21 days (range 15-30 days); (LT2 page N-16)

Routes – only oral ingestion (dermal and inhalation not considered)

Occurrence – although seven sources of data are cited, the main source of data was the ICR. Information in the survey included water quality parameters, such as turbidity and pH, along with process units in the plant and their sequences. The ICR survey is the most comprehensive database available for large systems. EPA identified an expected level of laboratory analytical method performance based on results with EPA Methods 1622/23 in the Information Collection Rule Supplemental Surveys, and also established the mean as the appropriate statistical measure to classify source water *Cryptosporidium* oocyst levels. The use of the arithmetic mean is advantageous for several reasons. The mean can be estimated more

Text Box 4. Exposure Analysis for the Long Term 2 (LT2) Enhanced Surface Water Treatment Rule (Source: U.S. EPA, 2006a).

(continued)
reliably than other statistical measures. For example, with a limited number of samples, the confidence interval around the mean is substantially narrower (i.e., less uncertain) than for a 90th percentile estimate. Defining a treatment trigger based upon a maximum value would be much less reliable than basing it on a computation involving multiple values, due to the uncertainty associated with any single sample measurement. The mean density also directly relates to the average risk of the exposed population and, therefore, provides a good measure for indicating relative risks from one site versus another (e.g., doubling the source water average density corresponds to about a doubling of the risk, assuming the same level of treatment at both sites). In contrast, the median would not be an informative or appropriate characterization because of the large numbers of non-detection measurements expected to occur, resulting in a large number of sites with median values equal to zero. The median would fail to distinguish differences between sites that had half or more of their measurements as zero and positive values for the remainder, and those that truly had measurements of zero. The input used to calculate infectivity included the following parameters: Density of oocysts = 0.0995 (range of 0.067 to 0.132 oocysts/L); percent infectious = 20% (range 15% to 25%)

Spatial temporal characterization – the ICR survey was conducted from 1997 through 1998. It consists of 18 months of data collected from all large systems in the United States serving over 100,000 people.

Other characteristics of the exposed populations – EPA considered as examples, Portland, ME, Portland, OR, Tacoma, WA, San Francisco, CA, and New York, NY populations, including the predicted number of people living with AIDS. EPA also considered small (< 10,000), medium (10,001 to 100,000), and large (>100,000) population size categories for analysis.

3.2.1. Identification of Media

In this document, identification of media refers to the specific water sources being considered in the risk assessment. Numerous water-related exposures can be considered using the tools, methods, and approaches described. For example, exposures through the ingestion of drinking water, ingestion of water during recreational activities (swimming), water reuse, inhalation of aerosolized biosolids particles, and ingestion of soil amended with biosolids are all exposure scenarios that are compatible with the tools in this MRA Tools document. To illustrate this point, MRAs for biosolids-related exposures have been conducted in the same manner as described herein for water-related exposures (Eisenberg et al., 2004, 2008; Gale, 2003, 2005), and MRAs for recycled wastewater in agricultural irrigation, swimming, and landscape irrigation practices have also been reported (Asano et al., 1992; Tanaka et al., 1998).

3.2.2. Routes of Exposure

The primary route of exposure[19] considered in a water-based MRA is usually ingestion, but can include other routes of exposures such as inhalation and dermal contact. Inhalation exposures may be significant for some microorganisms (spore-forming bacteria, *Legionella*, *Mycobacteria*, and

[19] Route of exposure refers to how the pathogen comes in contact with the vulnerable host receptor cells that support infection (e.g., inhalation, dermal contact, oral), whereas source of exposure refers to the physical matrix that carries the pathogen (e.g., air, water, food, soil).

some viruses). Similarly, dermal exposures (through intact skin or, more frequently, open cuts and scratches) may be important for some scenarios (e.g., *Pseudomonas aerugenosia*, *V. parahaemolyticus*, *V. vulnificus*). If data for a route of exposure are not available, it would not be possible to quantitatively evaluate within the risk analysis, however, it remains a major uncertainty and would need to be evaluated qualitatively in the risk characterization. For example, quantitative consideration of inhalation route would be possible for the enteric virus coxsackievirus, because respiratory pathway dose-response data are available (Couch et al., 1965).

Many water-based MRAs focus on exposure through drinking water and recreational activities, such as swimming and other activities, where ingestion of water is likely. However, there are also other potentially important routes of exposure that can also be of interest for which less data are currently available, such as exposures to recycled water, biosolids, as well as secondary and non-contact activities such as boating and fishing.

Routes of exposure should be discussed in the risk assessment documentation, including which routes are considered and which routes are not, as part of the scope of the risk assessment. Uncertainty in the exposure assessment should also be discussed for the routes that are not modeled or assessed.

3.2.3. Units of Exposure

The unit of exposure is generally a "dose", which is a specified number of pathogens that a person or population is exposed to. It is reasonable and convenient to consider exposure in terms of a magnitude, duration, and/or frequency. For example, the magnitude could be the number of plaque forming unites (PFU) per L, the duration and frequency could be 2 L per day for 360 days. For pathogens, exposure events are usually measured on a time scale of daily or shorter intervals. Historically, exposure time frames for pathogens have been based on the assumption that short-term (event-based) exposures are most relevant (e.g., per swimming event for recreational activities; per day for drinking water uses) rather than lifetime exposures. In contrast to chemical contaminants in water, the adverse health effects associated with human exposure to waterborne pathogens have been best documented for event-related (short-term, single exposure) rather than chronic exposure over extended periods of time. These short-term exposure timeframes have been used because infection requires that one or more pathogens be ingested and that at least one of the ingested pathogens establishes itself in or on cells somewhere within the GI tract of the host (Teunis and Havelaar, 2000). If no organisms have been ingested or none of those ingested succeed in passing all of the host barriers, infection does not occur. Note that short-term exposures do not necessarily imply that only short-term or minor adverse human health effects occur; for example, illnesses from some pathogens (e.g., *E. coli* O157:H7) can be severe and/or long-term or produce sequelae (Rangel et al., 2005).

Although there are chemicals for which a single exposure model is appropriate, such as teratogens that cause developmental defects or nitrate that can cause infantile methemoglobinemia, many cancer-causing chemicals exhibit increasing risk as duration of exposure lengthens (i.e., exposure over multiple years). This is because some chemicals can accumulate in the human body, and even for chemicals that can be purged from the body, the damage they cause may not be readily repairable. Therefore, damage may accumulate with each subsequent exposure. Although

accumulating damage is not necessarily an outcome of infection by pathogens, there are pathogens that generate toxins that behave similarly to chemicals in this respect. Therefore, it may be important to consider the mechanism by which pathogens cause illness symptoms when considering whether short-term, event-based exposures are the predominant relevant exposure pathway. Reinfection may also increase the potential for development of autoimmune disease. For example, the autoimmune disease reactive arthritis can be triggered by the following pathogens: *Chlamydia*, *Salmonella*, *Shigella*, *Yersinia*, *Brucella*, *Leptospira*, *Mycobacteria*, *Neisseria*, *Staphylococcus*, and *Streptococcus* (Girschick et al., 2008).

The exposure timeframe basic unit should be discussed within the context of the exposure scenarios. It may be difficult to determine if recurring exposure events are completely independent or not. For example, MRAs for drinking water commonly assume that all water consumed over the course of a single day is considered to be one dose, and consumption on subsequent days are considered to be independent events (U.S. EPA, 2006a). When exposures are considered to be completely independent (e.g., consumption on different days) the cumulative risk can be calculated as the result of independent repeated daily risk events. At the other end of the scale, when exposures are considered to be completely dependent, the doses can simply be added and treated as a single risk event (e.g., add total volume of water consumed through each serving of water over the course of a day). However, little data are available to describe the mechanisms of pathogen infection processes to support the assumption that all consumption within a specified period constitutes a single exposure event. Instead, the 24-hour timeframe is used for convenience and because it is a biologically reasonable timeframe for human digestive processes. The interdependence of exposure events may be important for some pathogens and may vary depending on characteristics of the host, the pathogen, and event specific conditions such as delivery matrix. Although exposure events may have varying interdependence, the assumption of independent exposure events as a default assumption is commonly used in MRA (Regli et al., 1991). Risk assessment tools for considering multiple exposures to a given pathogen are currently not sufficiently developed to recommend any specific tools.[20] The uncertainties associated with defining the unit of exposure should be discussed in the risk assessment.

Some examples illustrating how exposure can be computed are presented in Text Box 5 below. Note that various other uses of water, such as ingestion of crops irrigated with recycled water, and aquaculture, could have different units of exposure than those shown in Text Box 5. Soller et al. (2007) summarize the available literature and data that can be used for exposure to pathogens from three uses of recycled water, including body contact recreation, crop irrigation, and landscape irrigation. In all of those cases, exposure is through an ingestion route of exposure and is specified as volume ingested per day.

[20] Methods for considering exposure to multiple pathogens are also lacking.

Text Box 5. Examples of Time Units Associated with Exposure

- For recreational exposures, the time period associated with an exposure may be per swimming day or hours spent in the water. For example, risks from recreational exposures may be calculated using estimates of volume of ambient water incidentally (inadvertently) ingested over a given length of time (e.g., 50 mL/hour).
- For drinking water, the time unit of exposure is typically a single day, for example, 3 L per person-day (which represents the 95th percentile for consumption from EPA's Exposure Factors Handbook) (U.S. EPA, 2011).
- For shellfish, the associated exposure unit can be a meal or serving (without including how much constitutes a meal), number of shellfish consumed per meal, or weight of shellfish consumed per meal.

3.2.3.1. Volume Ingested

Quantitative data have been developed to characterize the volume of water that individuals and/or populations ingest through drinking water (U.S. EPA, 1997a, 2002d, 2006a, 2008, 2011) and recreational activities (Dufour et al., 2006; Schets et al., 2011). The data that are available for characterizing the volume of drinking water ingested has received the most attention and is by far the most comprehensive. The available data for characterizing the volume of water ingested during recreational activities derives from studies conducted in swimming pools. Data characterizing other routes of exposure are much more limited (Dorevitch et al., 2011). In fact, the limited data for assessing exposure (with the exception of that for drinking water) is currently an important limitation in MRAs.

Dufour et al. (2006) found that during swimming events lasting at least 45 minutes, children (≤18 years of age) ingested significantly more water (average = 37 milliliters [mL], range = 0 – 154 mL, N = 41) than adults (average = 16 mL, range = 0 – 53 mL, N = 12). The raw data provided by study authors (Dufour, personal communication) was fit to a statistical distribution. The best-fit distribution is lognormal with log mean (2.92) and log standard deviation (1.43). The median value of this distribution is 18.6 mL. Schets et al. (2011) also evaluated ingestion during swimming events. In their survey conducted in the Netherlands, 75,000 inhabitants representing the general Dutch population were asked to report the volume of water they swallowed as an estimated number of mouthfuls in four classes: (1) no water or only a few drops, (2) one to two mouthfuls, (3) three to five mouthfuls, and, (4) six to eight mouthfuls. Table 8 shows the self-reported results for the 8,000 questionnaires that were competed.

Table 8. Average Volume Water Swallowed (mL) per Swimming Event (Schets et al., 2011)

Setting	Age < 15 years old	Age ≥15 years old	
		Men	Women
Swimming pool	51 mL	34 mL	23 mL
Freshwater	37 mL	27 mL	18 mL
Seawater	31 mL	27 mL	18 mL

3.2.3.2. Temporal Nature of Exposure

The temporal nature of exposure includes the duration and frequency of the exposure. The duration is how long the exposure happens. For microbes duration is usually per event (e.g., swimming), per serving (e.g., food), or per day (e.g., drinking water consumed over a day). Frequency may be once, or compounded events over a year or a lifetime. For example The U.S. Census Bureau estimates 20 million (46 %) people (Age: 7-17 years) swim more than 5 times per year, and that 63 million (21%) people (Age: >17) swim more than 5 times per year (U.S. Census Bureau, 2011a, b).

Both endemic and episodic exposures are possible and MRA exposure assessments can be developed such that both of these exposure types are adequately modeled. For example, in a retrospective analysis of illnesses related to a water treatment system and distribution network, Westrell et al. (2003) estimated and compared illness rates for *Cryptosporidium*, rotavirus, and *Campylobacter* for customers of water from a conventional treatment plant using chemical disinfection to a hypothetical plant employing membrane filtration. Treatment processes were characterized by distributions (as opposed to use of non-parametric characterizations) and incidents sporadically included in the system model were (1) prolonged filter ripening periods every second day for one month each year (2) chlorine failure, (3) cross connections in the distribution system, and (4) pollution intrusion during low pressure transients. Monte Carlo simulation of the treatment and distribution process indicated that a major portion of potential infections occurred during normal plant and distribution system operation. Illnesses occurring at low rates and endemically may be difficult to detect and attribute to drinking water in public health surveillance. A counterpoint to the importance of endemic exposure is presented in an analysis of the 1993 Milwaukee cryptosporidiosis outbreak (Eisenberg et al., 2005). In that study, the combination of a drinking water treatment plant failure and ease of disease transmission led to a very large outbreak. The analyses demonstrate an apparent feedback between the production of infectious oocysts and the generation of unsafe drinking water. Here, the importance of the epidemic nature of the exposure is more important than endemic exposure. These examples illustrate the importance of considering and modeling both episodic and routine exposures and the potential importance of dynamic modeling as a component of MRAs.

3.2.4. Spatial Nature of Exposure

The spatial nature of the exposure for waterborne pathogens is suggested by the Clean Water Act designated use. For example, exposure during recreation (swimming, surfing, etc.) occurs while people are on or in the water and is geographically confined to where the water body is located. Exposure through drinking water is limited to the area that the public water supply serves, unless water is transported (e.g., transported by truck to a neighboring community to fill a special need). People may also move between water districts throughout the day as they travel for work or other reasons. Exposure through consumption of raw or partially cooked shellfish can also occur at locations removed from the water body from which the shellfish originated. Complex spatial distributions of exposure and the exposed people can make characterizing exposure patterns difficult; however, those patterns should be analyzed for their impact on the exposure assessment.

3.2.5. Behaviors of Exposed Population

Subpopulations can be defined by their susceptibility (e.g., intrinsic factors such as immune status or related factors) or by behaviors (extrinsic factors) that may cause them to be highly exposed (e.g., lifeguards, surfers, tri-athletes, other competitive swimmers versus casual bathers). In particular, plausible extreme behaviors should be noted, and the discussion should clarify to what degree individuals exhibiting those behaviors are addressed by the exposure scenario. Behaviors can also influence the routes of exposure and the spatial and temporal nature of exposure.

Specialized exposure scenarios, such as occupational exposures, can also be developed. This type of exposure consideration would most likely require that the risk assessment include both scientific and regulatory considerations. Risk assessments limited to occupational exposures to water that have caused infectious disease outbreaks are not common. However, there may be some occupations that have frequent water exposures in which a MRA may be of interest (e.g., lifeguards, wastewater treatment plant workers).

3.3. Exposure Profile and Linkage between Exposure Assessment and Other MRA Components

The exposure profile is a distillation of the most important information and analyses that are conducted during the exposure assessment. Each of the components of the exposure assessment describes the data and information that are available on that specific topic (i.e., occurrence, identification of media, units of exposure, routes of exposure, spatial and temporal nature of exposure, and characterization of exposed population). The exposure profile is a compilation summary of those data and analyses that will be used in conjunction with the human health assessment for estimating risk.

In the same manner as the problem formulation documentation, the formality of the exposure profile documentation can vary depending on the needs of the EPA Office conducting the assessment.

Consistent with the recommendations from the EPA-ILSI framework (ILSI, 2000) regarding the iterative and fluid nature of risk assessment, the exposure profile (as well as the host pathogen profile, as described in Section 4.3) should be critically evaluated by the risk assessors and managers to determine if the problem formulation component needs to be revisited and refined based on the availability of relevant data presented in the exposure profile. The linkage between the exposure assessment and the problem formulation phase of MRA is iterative.

3.3.1. Exposure Estimation

In the exposure profile the exposure estimate is presented. It serves as the critical linkage from the exposure assessment to the health effects assessment. Although the quantity and quality of data that will be available for any particular risk assessment will necessarily vary, the exposure profile provides critical input for the risk estimation.

The exposure estimate can include, as appropriate, a qualitative and/or quantitative evaluation of the magnitude, frequency, and patterns of exposure to a pathogen for each of the populations in each of the scenario(s) of interest. The exposure profile should also identify the specific assumptions that are made during the exposure assessment and uncertainties that are thought to be important for the risk assessment. Assumptions made during risk assessment are based on scientific judgment and consider the scope of the risk assessment as defined during problem formulation and planning and scoping.

3.3.2. Exposure Description

A description of the uncertainty associated with each element of the exposure assessment should be provided to the extent that it is reasonable and possible. Uncertainty analysis in drinking water MRAs has shown that exposure assessment can be a primary factor driving the output risk distributions. Thus, the description of uncertainty in the exposure profile is an important aspect of MRA.

The description of the assumptions and uncertainties related to exposure should be sufficient to provide an appropriate level of insight into the strengths and weaknesses of the assessment for evaluation during risk characterization. For example, Teunis and Havelaar (1999) used the exposure profile section of their *Cryptosporidium* in drinking water risk assessment to summarize the quantitative information on density of oocysts in raw water, recovery efficiency of the detection method, reduction by treatment, and amount of finished water that is consumed. The distribution type (e.g., negative binomial, beta, 2 choice binomial, log-normal) selected for each parameter as well as median and 95% range are presented in a table. A description of the Monte Carlo calculations and graphical as well as narrative discussion of the Monte Carlo simulation is also included. In this example, the exposure profile highlighted several important observations that affected the subsequent aspects of the MRA, including the following:

- Correction of oocysts counts for viability had little effect on the distribution of the density of oocysts in river water.
- The two distributions for river water and storage overlap, so that occasionally the treatment plant will be confronted with relatively high oocysts loads, even after passage of three reservoirs in series.
- Although treatment (physico-chemical) has a marked effect on oocysts densities (frequency distribution shifted by 4-logs), there is still a small probability of high densities of oocysts in treated water that is related to occasional reduced performance of the treatment plant.

As another example, Soller et al. (1999) used the exposure profile section of their drinking water risk assessment for rotavirus to summarize the exposure parameter assumptions. Below is a summary of the salient assumptions that were used in that MRA:

- The exposure model assumes that there is no upstream contamination or upstream contamination has been diluted to the point that the effects are negligible.
- The exposure model assumes that there are no animal (agricultural and grazing) sources of human infectious rotavirus.

- Acute-phase infected humans engaged in water recreation near the drinking water intake could be a significant source of rotavirus. However, this was not considered significant because site specific data that would be required to add this parameter is not available and body contact water recreation is likely to be insignificant during winter months, which is the time of year when rotavirus infections are most significant.
- Wastewater treatment plant effluent is the most important source of rotavirus and is assumed to have undergone secondary treatment with chlorine, contribute 5% of river volume, and contain 1 to 375 focus forming units/L of rotavirus.
- Rotavirus decay in source water results in 99% reduction between 3 and 30 days.
- Chlorine residual provides between 0 and 1.0-log reduction in rotavirus between the drinking water treatment facility and the tap.

4. Human Health Effects and Dose-Response

The human health effects assessment consists of the technical evaluation of data related to host characterization, evaluation of human health effects, and quantification of the dose-response relationship for contaminants in water. The problem formulation phase of the MRA precedes the human health effects assessment and may partially address many of the issues to be evaluated in the effects assessment, but the effects assessment is generally more detailed and quantitative than the problem formulation phase. The two components of the risk assessment, which may be conducted in parallel, are the characterization of exposure and the characterization of human health effects, including the dose-response relationship.

The output from the human health effects assessment is a host-pathogen profile that provides qualitative and or quantitative descriptions of the nature of the illness (e.g., morbidity and mortality statistics) and quantitative dose-response analyses for the scenario(s) developed during problem formulation. This chapter is divided into general health effects and dose-response modeling.

4.1. Human Health Effects Overview

The important health effects elements that can be considered during risk assessment are summarized below and discussed in the following sections, including (adapted from ILSI, 2000; see Table 1):

- duration of illness;
- severity of illness;
- morbidity, mortality, sequelae (long-term effects) of illness (including acute and chronic effects); and
- secondary transmission and immunity.

4.1.1. Duration of Illness

Duration of infectious disease illness is usually expressed in days. Duration can often be divided into duration of incubation (incubation period), duration of infection, duration of infectiousness (duration that host excretes the pathogen), and duration of disease symptoms. The scope of the risk assessment will determine the extent to which detailed information is required for each of these factors. If secondary transmission is expected to be significant, then the incubation period and duration of infectiousness may be important determinants of the magnitude of disease occurrence.

The incubation period for a disease is the interval from a person's exposure to the pathogen to the time they develop symptoms or clinical illness (or the period between the dose and some measurable response, such as shedding of the pathogen or serological response). Different diseases have different incubation periods, and this information can be used to help identify the pathogen responsible for a particular outbreak. Chronic sequelae from infections include all persistent and future effects on health (disability, recurrence of infection) and may extend for years after acute infection (see Section 4.1.3 below). A brief summary of some incubation periods for several waterborne diseases is presented in Table 9.

Table 9. Typical Incubation Periods for Some Waterborne Pathogens

Pathogen	Incubation Period and Reference[a]
Cryptosporidium parvum	2 to10 days (average 7 days)[b]
Giardia lamblia	1 to 2 weeks (average 7 days)[c]
Shigella spp.	16 to 72 hours[d]
Campylobacter jejuni	2 to 5 days (Trachoo, 2003)
Escherichia coli O157:H7	1 to 8 days (average 3 to 4 days) (Weir, 2000)
Norovirus	24 to 48 hours (APHA, 2004)
Rotaviruses	< 4 days (average 1 to 2 days) (Aitken and Jeffries, 2001)

[a] Defined as time from exposure to onset of first symptoms.
[b] http://www.cdc.gov/ncidod/dpd/parasites/cryptosporidiosis/factsht_cryptosporidiosis.htm.
[c] http://www.cdc.gov/ncidod/dpd/parasites/giardiasis/default.htm.
[d] http://pathport.vbi.vt.edu/pathinfo/pathogens/Shigella.html.

4.1.2. Severity of Illness

The severity of illness, morbidity, mortality, and chronic sequelae of illness are all factors that need to be considered in the choice of health endpoints considered in the risk assessment. Severity of illness is often difficult to quantify because disease symptoms often include subjective descriptions. Severity of illness can be measured by more objective parameters, such as T-cell count or other biological markers (e.g., liver function). Number of physician visits, hospitalizations, or emergency room visits may also be used to assess severity, but these measures have the disadvantage that they depend on the availability of such services, and cultural and social values related to the use of medical services, and costs. Severity of infection does not necessarily equate to severity of illness. Individuals that are infected and are able to transmit the disease, but do not exhibit symptoms, are known as asymptomatic carriers. The length of time an individual remains in the carrier state can vary based on pathogen and host factors. Severity of illness and severity of infection are usually used in reference to an individual (as opposed to a population). An individual may also exhibit varying degree of infectiousness during the course of an infection and infectiousness between individuals can be different. Severe illness may or may not be accompanied by severe infectiousness.

4.1.3. Morbidity, Mortality, and Sequelae

Morbidity and mortality measures can also be used to characterize disease burdens within a population. Morbidity is a measure of the proportion of people who are afflicted with a given disease or who display a given symptom per unit of population (e.g., per 1000 people, per 100,000 people). Mortality is a measure of the number of deaths per unit population, or number of deaths out of the diseased population. Both morbidity and mortality are most commonly expressed as annual rates (or rates during an outbreak).

Sequelae of illness, which are more commonly referred to as "chronic sequelae," are conditions that occur after infection has occurred. Because chronic symptoms may be removed in time from the acute infection, it is often harder to demonstrate a correlation between infection and symptoms.

Furthermore, the type of epidemiological study design that can detect chronic sequelae (i.e., retrospective cohort study design) is not commonly conducted for waterborne illnesses. Section 4.1.4 includes information on chronic sequelae from infection with several waterborne pathogens of public health significance.

U.S. Centers for Disease Control and Prevention (CDC) periodically reports estimates of the incidence of foodborne illness calculated from total estimated illness (Mead et al., 1999; Scallan et al., 2011a,b) and reports on waterborne disease surveillance (CDC, 1993, 1996, 1998b, 2000, 2002, 2004b, 2006b, 2008). These resources are helpful for framing the human health effects of pathogens.

4.1.3.1. Acute and Chronic Health Effects of Microbial Contaminants

Waterborne microbial contaminants cause a range of acute health effects, including but not limited to gastrointestinal illness, eye infections, ear infections, respiratory infections and other illnesses described below for a set of pathogens. Symptoms such as bloating, lethargy, general malaise, aches, malabsorption, weight loss, anorexia, and dehydration can result following infections with *Giardia and Cryptosporidium*. Other symptoms include dysentery (characteristic of *Vibrio* species); influenza-like symptoms such as fever, chills, headache, myalgia (e.g., *Legionella*); hepatitis (Hepatitis A and E); and meningitis, which has been associated with *Enterovirus*, *Aeromonas hydrophila*, *Campylobacter*, and *Yersenia enterocolitica.* Chronic diarrhea is typically associated with enteric protozoa such as *Cryptosporidium* and *Giardia*, but some bacteria, such as some adherent enteropathogenic *E. coli,* can also result in this condition. Infection with pathogenic *E. coli* can result in kidney disease, and hypertension. People who are immunocompromised are more likely to experience chronic diarrhea. Chronic sequelae (long-term health effects) may also develop following infections from waterborne infections. Overviews of the public health significance, including chronic sequelae, of several waterborne viruses, bacteria, amoebae, and protozoa are provided below. Other resources for pathogen information include, a review of waterborne pathogens (Craun et al., 2010), the American Society for Microbiology's (ASM) Manual of Clinical Microbiology (ASM, 2011), and FDA's "Bad Bug Book" (FDA, 2006).

VIRUSES

Adenovirus

Adenoviruses infections most commonly cause lower and upper respiratory disease, AGI (especially due to serotypes 40 and 41), acute conjunctivitis, pharyngoconjunctival fever, and urinary tract infections (Enriquez and Thurston-Enriquez, 2006). Children, infants, the elderly, immunocompromised, and other sensitive populations can be especially affected (Jiang, 2006). Infection is most often spread via direct contact with an infected individual, fecal-oral route, or recreational water activities. Although not thought to be the most common route of transmission, there have been two drinking water outbreaks reported in Europe in which enteric adenoviruses may have been a cause of acute gastroenteritis (Mena and Gerba, 2009). Adenoviruses in drinking water may also contribute to viral infections of unknown etiology (Ko et al., 2003).

Neurological sequelae from Reye's syndrome following infections by organisms such as adenovirus have been reported. For example, studies of Reye's Syndrome survivors have shown

sequelae ranged from severe psychomotor retardation to mild specific perceptual and/or language impairments, IQ deficits, and chronic behavioral deficits that lasted from 6 to 18 months after infection (Davidson et al., 1978; Brunner et al., 1979; Shaywitz et al., 1982).

Astrovirus

Astroviruses are transmitted by person-to-person contact via the fecal-oral route, and especially via the ingestion of contaminated food or water or contact with fomites (Abad et al., 1997, 2001; Schwab, 2006). Globally, much of the endemic- and outbreak-related cases of gastroenteritis in children are caused by astroviruses (Walter and Mitchell, 2000; Glass et al., 2001). Infection with astroviruses characteristically results in GI illness, most commonly self-limiting diarrhea, vomiting, and mild dehydration. However, more severe cases are possible. Astroviruses continue to be linked to documented waterborne disease outbreaks both in the United States and abroad (e.g., Guix et al., 2005; Smith et al., 2006).

There are limited data on the occurrence of sequelae following astrovirus infection. Chronic diarrhea has been reported lasting over 1 month (Grohmann et al., 1993; Lin et al., 2008).

Calicivirus (including Norovirus)

Caliciviruses, which include norovirus, are highly contagious and transmitted primarily through the fecal-oral route—most commonly by consumption of fecally contaminated food or water or through person-to-person spread (Schwab and Hurst, 2006; CDC, 2009b). However, environmental and fomite contamination may also act as an important source of transmission and infection. Waterborne outbreaks of norovirus disease in community settings are often caused by sewage contamination of wells and recreational water (CDC, 2009b) and are widely considered to be the causative agents in many viral waterborne outbreaks of unknown etiology (Maunula et al., 2005; U.S. EPA, 2006c). Although pre-symptomatic viral shedding may occur, shedding usually begins with onset of symptoms and may continue for up to 2 weeks after recovery from illness (CDC, 2009b). The incubation period for norovirus-associated AGI in humans is usually between 24 and 48 hours (median in outbreaks 33 to 36 hours), but cases can occur within as little as 12 hours of exposure.

Infection usually presents as acute-onset vomiting, watery non-bloody diarrhea with abdominal cramps, and nausea. Low-grade fever may also occur, though dehydration is the most common complication, especially among the young and elderly. Symptoms typically last 24 to 60 hours, and patients recover completely. Although asymptomatic infection may occur in as many as 30% of infections, it role in calicivirus/norovirus transmission remains poorly understood.

Enterovirus

Enteroviruses are associated with a variety of clinical syndromes, ranging from mild and self-limiting to severe, potentially fatal conditions. Non-poliovirus enteroviruses likely cause 10 to 15 million symptomatic infections and 30,000 to 50,000 hospitalizations in the United States each year (Kim et al., 2001; CDC, 2006a). Enterovirus infections can cause a range of illnesses including AGI (fever, nausea, diarrhea, or vomiting), conjunctivitis, skin rashes, cold and flu-like illnesses, muscle inflammation, arthritis, paralysis, respiratory disease, meningitis, myocarditis, organ failure, and even death. Infection is most often spread via the fecal-oral route, contact with

feces or virus-contaminated objects, direct contact with an infected individual, or recreational water activities. Enteroviruses were responsible for at least one waterborne disease outbreak in the developed world (Borchardt et al., 2003) and have been detected in U.S. source waters and treated drinking water (Borchardt et al., 2004; Vivier et al., 2004). Enteroviruses in drinking water may also contribute to viral infections of unknown etiology and unknown levels of endemic disease (U.S. EPA, 2006c).

Enteroviruses potentially play a role in the development of chronic diseases, such as juvenile diabetes and chronic fatigue syndrome (Behan and Bakheit, 1991; Fohlman and Friman, 1993), myalgic encephalomyelitis (Lloyd et al., 1988), and myocarditis (Kim et al., 2001). Some patients who have paralysis or encephalitis do not fully recover, and persons who develop heart failure from myocarditis require long-term care for their conditions (CDC, 2010).

Hepatitis A Virus

Hepatitis A virus caused an estimated 32,000 infections in the United States in 2006, although numbers of infections decrease each year, presumably due to the increased vaccination of children (CDC, 2009c). Only rarely fatal, hepatitis A infections typically cause acute liver disease lasting a few weeks to several months. Symptoms include fever, tiredness, nausea, decreased appetite, and abdominal discomfort, followed by jaundice (Sobsey, 2006). The disease is more severe in adults than it is in children. In the United States, Hepatitis A infection is most often spread from person-to-person (often within the household) via the fecal-oral route and contact with feces or virus-contaminated objects, but also by ingesting contaminated food or water, or through recreational water activities. Drinking-water related Hepatitis A outbreaks have occurred in the United States (e.g., Hejkal et al., 1982; U.S. EPA, 2006c; Yoder et al., 2008).

Hepatitis A has been associated with arthritis (Willner et al., 1998; Fan et al., 2009), specifically mimicking autoimmune disease (Sridharan et al., 2000).

Hepatitis E Virus

Hepatitis E virus infections typically cause acute liver disease lasting a couple weeks to a few months (Gerba, 2006). Although rarely fatal for most individuals (the most commonly affected age group is young to middle-aged adults), the disease has a high mortality rate (15 to 20%) in pregnant women (Mushahwar, 2008). Major waterborne Hepatitis E virus epidemics have been documented in developing countries, but the disease is rare in the United States (Arguin et al., 2008). Hepatitis E virus is transmitted via the fecal-oral route and infection is spread primarily through fecally-contaminated drinking water or food, such as shellfish (WHO, 2009). Human strains of Hepatitis E virus have experimentally infected pigs, and porcine strains have experimentally infected primates (WHO, 2004; Halbur et al., 2001).

Rotavirus

Rotavirus infection is a leading cause of acute severe gastroenteritis in young children and infants and worldwide (Abbaszadegan, 2006). They are excreted in very large quantities in the feces of infected persons, so are present in relatively high densities in wastewater and environmental water (Ansari et al., 1991). Each year, rotavirus infections account for an estimated 55,000 to 70,000 hospitalizations in the United States (CDC, 2008). Symptoms of rotavirus infection include

diarrhea, fever, and vomiting, leading to dehydration—especially in young children and infants. Since the introduction of vaccines against rotavirus in 2006, numbers of cases in the United States have declined. Rotavirus is highly contagious and the most infectious of the enteric viruses (Gerba et al., 1996b). Infection is most often spread via the fecal-oral route, contact with feces or virus-contaminated objects, direct contact with an infected individual, or from contaminated food or water. Although not common, drinking-water related rotavirus outbreaks have occurred in the United States (e.g., Hopkins et al., 1984, 1985). Rotaviruses have been detected in treated drinking water (Abbaszadegan et al., 1999) and nondisinfected well water (Borchardt et al., 2003). Rotaviruses in drinking water can also contribute to waterborne outbreaks of unknown etiology and unknown levels of endemic disease.

BACTERIA

Aeromonas hydrophila

Aeromonads can cause mild to severe GI illness following consumption of contaminated food or water and can vary in severity from self-limiting watery diarrhea with or without fever to an acute, cholera-like illness with profuse watery diarrhea (Moyer, 2006). Severe infections most commonly occur in young children and immunocompromised persons. *Aeromonas hydrophila* and related aeromonads can also cause a variety of soft-tissue infections within hours following dermal penetration injuries, leading to fever, pain, swelling, erythema, and edema. A wide range of potentially serious complications can follow soft-tissue infection including cellulitis, arthritis, sepsis, meningitis, and pneumonia (Noonburg, 2005). Other complications associated with *Aeromonas* infection include hemolytic uremic syndrome, septicemia, meningitis, peritonitis, wound infections, respiratory tract infections, and ocular infections (U.S. EPA, 2006d).

Arcobacter butzleri

Infections from *A. butzleri*, which was not distinguished from *Campylobacter* until recently (Snelling et al., 2006), can result from fecal-oral transmission and can cause AGI, including diarrhea associated with abdominal pain, nausea, vomiting and fever, and which can become life-threatening in immunocompromised persons and populations. It is considered to be an emerging foodborne pathogen of major significance in humans (Lehner et al., 2005; Snelling et al., 2006; Cervenka, 2007). Although little is known about the mechanisms of pathogenicity or potential virulence factors of *Arcobacter* spp., there is increasing evidence that livestock animals may be a significant reservoir and that it may be zoonotic.

Campylobacter jejuni

The most important *Campylobacter* species from a public health perspective is *C. jejuni*, which is also a major zoonotic pathogen. The most common health effect of *C. jejuni* infection is acute self-limited AGI, characterized by diarrhea, fever, and abdominal cramps (Butzler, 2004). The incubation period is typically 2 to 5 days but can extend up to 10 days. In about 50% of patients, diarrhea is preceded by a febrile period associated with malaise, myalgia, and abdominal pain; fresh blood may also appear in the stools by the third day. Most *C. jejuni* infections are sporadic (often outbreak-related) in nature, and in most cases, the source of infection is never determined (Hänninen et al., 2003). In drinking water outbreaks attributed to *C. jejuni* (e.g., Jones and Roworth, 1996; Holme, 2003) the drinking water source is usually shown to be fecally-

contaminated by runoff of surface water after rain or by leakage of a sewage pipe close to the drinking water pipeline.

Reactive arthritis is a rare complication of infection, and is strongly associated with people who have the human lymphocyte antigen B27 (FDA, 2009). The arthritis typically occurs within 2 weeks after the onset of GI illness, but onset time can range from 4 to 35 days (Anonymous, 1998). The condition is characterized by infection at a distant site, whereby joint inflammation occurs without typical evidence of sepsis at the affected joint(s) (Shirtliff and Mader, 2002).

Reiter's syndrome is an aseptic arthritis (Kim et al., 2007) that is triggered by an infectious agent located outside the joint. It also has a strong association with human leukocyte antigen-B27. Pope et al. (2007) reviewed the epidemiological literature on *Campylobacter*-associated reactive arthritis and found that follow-up for long-term sequelae related to *Campylobacter* and Reiter's syndrome was largely unknown.

Campylobacter jejuni is increasingly recognized as a risk factor for Guillain-Barré syndrome, a common cause of neuromuscular dysfunction (Bunning et al., 1997) that occurs in one in 1,000 infections (Butzler, 2004). Guillain-Barré syndrome is the most serious complication of *Campylobacter* infection (Kaldor and Speed, 1984; Nachamkin et al., 2000; Yuki, 2001; Ang et al., 2002; Gilbert et al., 2004). Studies by McCarthy and Giesecke (2001) have shown that the risk of developing Guillain-Barré syndrome during the 2 months following a symptomatic episode of *C. jejuni* infection is approximately 100× higher than the risk in the general population.

Immunoproliferative small intestinal disease is an infection associated lymphoma disease (Al-Saleem and Al-Mondhiry, 2005). Molecular and immunohistochemical studies have demonstrated an association with *C. jejuni* (Lecuit et al., 2004). This disease involves the proximal small intestine resulting in malabsorption, diarrhea, and abdominal pain**,** as well as weight loss, intestinal obstructions, and abdominal masses (Gilinsky et al., 1987; Rambaud et al., 1990; Fine and Stone, 1999; Al-Saleem and Al-Mondhiry, 2005). Symptoms are often chronic, and patients may experience mild symptoms for 5 to 10 years before developing higher-grade illness or lymphoplasmcytic and immunoblastic lymphomas (Al-Saleem and Al-Mondhiry, 2005).

Escherichia coli O157

E. coli O157 has been found worldwide in fecally contaminated surface water, groundwater, non-chlorinated or inadequately treated drinking water and swimming pools, soil and sediment, and a wide variety of foods (see review by Muniesa et al., 2006). It has been documented to cause both endemic and outbreaks of illness in humans through fecal-oral transmission, consumption of fecally-contaminated food and water, and contact with infected persons and animals (Rangel et al., 2005). The clinical symptoms of infection vary from non-bloody diarrhea to bloody diarrhea, referred to as hemorrhagic colitis. Hemorrhagic colitis is a serious, life-threatening condition and can lead to hemolytic uremic syndrome, which results in renal damage and possibly death.

Enterohemorrhagic *E. coli* infections can lead to long-term or permanent kidney damage and renal disease. Persons who develop chronic kidney failure may require lifelong dialysis or a kidney transplant. Garg et al. (2003) conducted a comprehensive review and meta-analysis of the current literature along with expert consultations in order to quantify the long-term renal prognosis of

patients with diarrhea-associated hemolytic uremic syndrome. A higher severity of acute illness was strongly associated with worse long-term prognosis. Studies with a higher proportion of patients with central nervous system symptoms (coma, seizures, or stroke), had a higher proportion of patients who died or developed permanent end stage renal disease at follow-up. Death or end stage renal disease occurs in about 12% of patients with diarrhea-associated hemolytic uremic syndrome, and 25% of survivors demonstrate long-term renal sequelae.

E. coli is also implicated in rheumatoid diseases (Locht and Krogfelt, 2002; Schiellerup et al., 2008), including reactive arthritis and Reiter syndrome with similar pathologies as described above.

Helicobacter pylori

This common infectious pathogen occurs worldwide and is associated with a variety of upper GI conditions (Bellack et al., 2006), including gastritis and peptic and duodenal ulcer disease. At present, the route of transmission to susceptible hosts remains largely unknown, although fecal-oral and person-to-person contact have been proposed as possible routes of exposure (Brown, 2000; Bellack et al., 2006). In addition, fecally-contaminated drinking water may also be a source of transmission and subsequent infection (McDaniels et al., 2005; Reavis, 2005; Azevedo et al., 2006).

H. pylori causes chronic gastritis (Chey and Wong, 2007), with a small proportion of affected patients (6 to 20%) proceeding to more severe clinical disease (Parsonnet et al., 1991), and can lead to gastric cancer (Wang et al., 2007). It has been estimated that *H. pylori*-positive patients have a 10 to 20% lifetime risk of developing ulcer disease and only a 1 to 2% risk of developing gastric cancer (Kuipers et al., 1995; Kuipers, 1999; Ernst and Gold, 2000). *H. pylori* is classified as a Class I carcinogen by the International Agency of Research on Cancer (IARC, 1994). According to the CDC, over 90% of duodenal cancers, and 80% of gastric ulcers are caused by *H. pylori* (CDC, 1998a).

Legionella pneumophila

Legionella pneumophila can cause respiratory disease in humans when a susceptible host inhales aerosolized water containing the bacteria or aspirates water containing the bacteria (Fields et al., 2002; Hall, 2006). Legionellosis classically presents as two distinct clinical entities, (1) Legionnaires' disease, a severe multisystem pneumonia-like disease that includes fever, nonproductive cough, headache, myalgias, rigors, dyspnea, diarrhea, and delirium (Fraser et al., 1977); and (2) Pontiac fever, a self-limited flulike illness (Glick et al., 1978). However, many persons who seroconvert to *Legionella* will remain entirely asymptomatic (Boshuizen et al., 2001). It has been estimated that 8,000 to 18,000 persons are hospitalized with legionellosis annually in the United States (Fields et al., 2002). The disease is a major concern of public health professionals, and especially organizations and individuals involved with maintaining building water systems. However, legionellosis is generally considered to be a preventable illness because controlling or eliminating the bacterium in water reservoirs will typically prevent the disease.

Legionella infections are associated with long-term development of chronic inflammatory and fibrotic reactions (pulmonary fibrosis) in the human lung (U.S. EPA, 1999b).

Mycobacterium avium

People can be exposed to *M. avium* in water through drinking, swimming, and bathing activities and through ingestion or inhalation of water vapor or droplets (Primm et al., 2004). As opportunistic waterborne pathogens, exposure to *M. avium* can infect the lungs and lead to cough, fatigue, weight loss, fever, and night sweats, especially in immunocompromised persons and populations (e.g., AIDS patients) (Primm et al., 2004; LeChevallier, 2006). Disseminated MAC infections in individuals with AIDS accounts for the majority of MAC-related morbidity and mortality in the United States.

MAC has also been associated with hypersensitivity pneumonitis (Lacasse et al., 2003; Fink et al., 2005; Hanak et al., 2007). Symptoms may be acute with flu-like symptoms, fever, chills, malaise, headache, and cough, or chronic, characterized by dyspnea (a cough that may be dry or productive) and weight loss (Lacasse et al., 2003; Field et al., 2004).

Plesiomonas shigelloides

Infection by *P. shigelloides* most commonly results from ingestion of fecally-contaminated food (especially seafood) and water—especially in immune-compromised persons (Wong et al., 2000; CDC, 2009a). Extra-intestinal health effects, most commonly observed in sensitive populations, include septicemia, meningitis in neonates, cellulitis, septic arthritis, and acute cholecystitis. However, in immunocompetent persons, infected persons typically develop self-limiting watery diarrhea that can last for 2 weeks or more (Huq and Islam, 1983; Wong et al., 2000; CDC, 2009a).

Salmonella enterica

Most human infections can be traced back to contaminated food products, although water-related outbreaks, including drinking water outbreaks, continue to be reported (Schuster et al., 2005; CDC/MMWR, 2006; Covert and Meckes, 2006; Craun et al., 2006). The incubation period ranges from 18 to 48 hours after ingestion and illness is usually characterized by acute, self-limiting AGI (though some infections can be severe), fever, and septicemia, that lasts from 2 to 5 days (Covert, 1999).

A broad range of chronic sequelae have been reported, such as reactive arthritis, Reiter's syndrome, rheumatoid syndromes, pancreatitis, osteomyelitis, myocarditis, colitis, choleocystitis, and meningitis have been reported as a consequence of *Salmonella* infection (Thomson et al., 1995; Dworkin et al., 2001; Motarjemi, 2002).

Shigella sonnei

Shigellosis, commonly known as acute bacillary dysentery, is characterized by the passage of loose stools mixed with blood and mucous and accompanied by fever, abdominal cramps, and tenesmus (Sur et al., 2004; Moyer and Degnan, 2006). The incubation period of shigellosis is typically 1 to 4 days, which is usually followed by sudden onset of AGI symptoms. In mild cases, the disease may be self-limiting, but severe disease requires appropriate medication. The disease is communicable as long as an infected person continues to excrete the organisms in feces at high levels, which can persist up to 4 weeks from the onset of illness. Thus, humans are the principal reservoir of infection (Sur et al., 2004), and transmission usually occurs via fecally-contaminated water and food or through person-to-person contact (Niyogi, 2005).

In addition to acute and chronic GI effects, shigellosis can cause central nervous disease such as seizures and convulsions ((Barrett-Connor and Connor, 1970; Khan et al., 1999). *Shigella* has also been associated with urinary tract infections (Ekwall et al., 1984). Other complications from shigellosis include bacteremia, hemolytic uremic syndrome, toxic megacolon (reviewed by Gupta et al., 2004), and encephalopathy (Goren et al., 1992; Ferrera et al., 1996). Both reactive arthritis and Reiter's syndrome are documented long-term sequelae from shigellosis (Gupta et al., 2004), and the pathologies of these conditions are similar to those of other enteric pathogens such as *Salmonella*.

Vibrio cholerae

V. cholerae is the cause of cholera, a GI illness of global and historic importance. In nature, *V. cholera* is often associated with protozoa (including *Acanthamoeba*), zooplankton, sediments, and shellfish, and can become a normal component of the aquatic microbiota (especially saline waters) is many parts of the world, including the United States. Although contaminated food remains the predominant source of infection and illness, contaminated water remains a highly important source of *V. cholerae* infections worldwide (Toranzos et al., 2006). Infections have been shown to vary from completely asymptomatic to very severe, profuse watery diarrhea and vomiting that can lead to rapid dehydration and death within one to five days. No outbreaks of cholera have been reported in U.S. drinking water supplies in several decades (see Steinberg et al., 2001).

Yersinia enterocolitica

Most human infections result from fecal-oral transmission and can usually be traced to contaminated food or water. *Y. enterocolitica* can cause a variety of symptoms depending on the age of the person infected, but most commonly GI, fever, and sometimes vomiting in children that is typically self-limiting and can last for up to 2 weeks (Fricker, 2006). However, post-infectious chronic sequelae include a form of reactive arthritis and a pseudoappendicitis syndrome (Sharma et al., 2003).

Yersinia elicits rheumatoid disease, similar to many other enteric bacteria such as *Shigella*, *Campylobacter*, and *E. coli*. Post-infectious sequelae of *Yersinia* infection include a form of reactive arthritis and joint symptoms (Wolf et al., 1991; Luo et al., 1994; Sharma et al., 2003; Schiellerup et al., 2008).

AMOEBAS

Acanthamoeba

Acanthamoeba spp. are opportunistic pathogens that produce granulomatous amebic encephalitis—a chronic central nervous system disease of immunocompromised hosts—and various other diseases, including keratitis and pneumonitis (Marshall et al., 1997; Visvesvara and Moura, 2006). To date, keratitis is the only water-related syndrome caused by *Acanthamoeba*. *Acanthamoeba* was determined to be the causative agent of keratitis in an U.S. outbreak involving a contaminated municipal water supply (Meier et al., 1998; Karanis et al., 2007). Untreated keratitis can lead to loss of visual acuity and blindness (Illingworth and Cook, 1998; Marciano-Cabral et al., 2000; Marciano-Cabral and Cabral, 2003; Khan, 2006; Awwad et al., 2007;

Thebpatiphat et al., 2007; Visvesvara et al., 2007). The infection is also known to cause chronic encephalitis in immune-deficient individuals (U.S. EPA, 2003h).

Blastocystis hominis

B. hominis is transmitted via the fecal-oral route through ingestion of contaminated food and water (Garcia, 2006a). When present in large numbers, *B. hominis* can cause GI illness, including diarrhea, cramps, nausea, fever, vomiting, and abdominal pain that may require medical attention. Although the severity and duration of these symptoms can be increased in immunocompromised persons, in healthy adults their presence may be asymptomatic (Chen et al., 2003). Globally, *B. hominis* has been attributed to four and possibly five documented waterborne disease outbreaks (Karanis et al., 2007)—including a 2005 gastroenteritis outbreak involving an inadequately disinfected public drinking water system in Turkey (Tuncay et al., 2008).

Entamoeba histolytica

Transmission is via the fecal-oral route, directly from person-to-person contact, or through contaminated food, water, or fomites (Marshall et al., 1997; Keene, 2006). In this manner, persons become infected and can become ill after ingesting cysts. The incubation period is highly variable, ranging from a few days to several months. Although most *E. histolytica* infections are asymptomatic, amebiasis can be a serious and even life-threatening disease, especially in developing nations. It can also result in recurrent diarrhea of varying severity, often with blood or mucous, fever, abdominal pain, and tenesmus. While drinking water-related outbreaks have been documented outside of the United States (e.g., Chen et al., 2001; Karanis et al., 2007), there have been no reports on the occurrence of *E. histolytica* in source or finished drinking water, or U.S. waterborne disease outbreaks since 1980 (Marshall et al., 1997).

Naegleria fowleri

Primary amebic meningoencephalitis associated with exposure to *Naegleria fowleri* is a very rare but usually fatal central nervous system illness that has led to several deaths in the United States (e.g., Barnett et al., 1996; Gyori, 2003; Marciano-Cabral et al., 2003; Visvesvara and Moura, 2006). Once entering into the nostrils of swimmers and others engaging in water sports, *N. fowleri* penetrates the mucosal layer and migrates along the olfactory nerve tracts and eventually reaches the brain (Schuster and Visvesvara, 2004). The often fatal disease is typically acquired by immune-competent children and young adults while swimming and diving in untreated warm freshwater lakes and ponds during summer months (Barnett et al., 1996; Marciano-Cabral et al., 2003; Schuster and Visvesvara, 2004).

PROTOZOA

Cryptosporidium

Cryptosporidium is a zoonotic waterborne pathogen of global public health importance. The CDC estimates that there are 700,000 cases of *Cryptosporidium* infections a year in the United States (Scallan et al., 2011a). Worldwide, the estimated number of annual cases of cryptosporidiosis exceeds several million and has been documented in almost 100 countries and on every continent except Antarctica (Casemore et al., 1997; Fayer et al., 1997; Cacciò, 2005). *Cryptosporidium* is a well-known cause of opportunistic infections among AIDS patients (Roy et al., 2004) and a

common cause of outbreaks of GI illness (CDC, 2008). Surveillance data indicate that infections are common among immunocompetent individuals and that 90% of cases are not involved in outbreaks (Dietz et al., 2000). Valid species of *Cryptosporidium* have been described in more than 155 mammalian species, >30 avian species, 57 reptilian species, 9 species of fish, and 2 amphibian species (O'Donoghue, 1995; Fayer, 2004).

Most clinical cases of cryptosporidiosis involve infection by *C. parvum* or *C. hominis*. The most common feature of cryptosporidiosis is profuse, watery diarrhea. Other clinical signs of infection include dehydration, fever, anorexia, weight loss, weakness, and progressive loss of condition (O'Donoghue, 1995). Recovery is usually spontaneous within 1 to 2 weeks of infection. Developmental stages of the parasite are often seen within the small intestine and occasionally elsewhere (stomach, colon, liver, lungs). In general, the development of cryptosporidiosis depends on the species, age, and immune status of the host (Fayer et al., 1997). Younger persons with less developed or compromised immune systems are generally more susceptible to severe infection than healthy adults (O'Donoghue, 1995).

Several Human volunteer feeding studies have been conducted to determine the infectivity of *C. parvum* and *C. hominis* in healthy adults in order to predict the likelihood of enteric infection following exposure to contaminated drinking water (DuPont et al., 1995; Chappell et al., 1996; Okhuysen et al., 1999, 2002). A summary of the dose-response data for all six of the tested isolates has been published by EPA (U.S. EPA, 2006b).

Cyclospora

Endemic in much of the developing world, but most common in tropical and subtropical areas, cyclosporiasis is considered an emerging GI disease in developed nations. It has been identified as the cause of several outbreaks in North America and Europe and with traveler's diarrhea (Herwaldt, 2000; Karanja et al., 2007). Cyclosporiasis transmission has primarily been linked to fecally-contaminated foods and water and is associated with a wide variety of GI symptoms, such as loose or watery diarrhea, nausea, vomiting, abdominal cramps, or appetite loss. It can also be associated with fever; chills; muscle, joint, or generalized body aches; headache; or fatigue. However, the exact mechanism of disease transmission remains unknown because freshly excreted oocysts are not infective and require days to weeks to mature sufficiently to become infective upon consumption; thus, direct person-to-person transmission of the disease is unlikely (Herwaldt, 2000). Although asymptomatic infections are known to occur, especially in immunocompetent children, the onset of symptoms in naïve populations observed in outbreaks is typically 1 to 14 days after exposure and is often accompanied by a characteristic waxing and waning of symptoms over time. In endemic countries, symptoms begin approximately 5 to 8 days after exposure and may persist for over a month, although watery diarrhea remains the most common health outcome.

Reports of *C. cayetanensis* sequelae such as biliary disease, aculculous cholecystitis, Guillain-Barré syndrome and reactive arthritis syndrome following prolonged infection were found in the literature (Sifuentes-Osornio et al., 1995; Richardson et al., 1998; Connor et al., 2001).

Giardia

Giardia intestinalis (also known as *G. lamblia* and *G. duodenalis*) is a zoonotic waterborne pathogen of global public health concern. It is the most common intestinal parasite identified by

public health laboratories in the United States (Rose et al., 1991; Kappus et al., 1994). CDC estimates that there are approximately 1.2 million to 2 million illnesses annually in the United States due to *Giardia* (Mead et al., 1999; Scallan et al., 2011a). Thirty percent of the U.S. general population has seropositivity for *Giardia* (indicates current or past infection) and 7% of small children are asymptomatically infected with *Giardia* (Frost and Craun, 1998). More than 100 waterborne giardiasis outbreaks have been reported worldwide from the beginning of the previous century till 2004 (Plutzer et al., 2010). *Giardia* also infects a wide variety of domestic and wild mammals (e.g., cats, dogs, cattle, deer, and beavers) (Thompson, 2000; Ballweber et al., 2010).

High risk groups for giardiasis include infants and young children, travelers to developing countries, the immunocompromised, and persons who consume untreated water from lakes, streams, and shallow wells (U.S. EPA 1998a; CDC, 2008). A wide spectrum of symptoms are associated with giardiasis, which range from asymptomatic infection and acute self-limiting AGI to persistent chronic diarrhea, which sometimes fails to respond to treatment. Asymptomatic infection is very common (50 to 75% of infected persons are symptomatic) (Mintz et al., 1993; U.S. EPA 1998a). Symptoms of giardiasis include diarrhea, abdominal cramps, bloating, weight loss, and malabsorption (Rodrigquez-Hernandez et al., 1996; Hellard et al., 2000; Thompson, 2000). Case reports also indicate that giardiasis might be associated with the development of reactive arthritis (Tupchong et al., 1999). *Giardia* infection is frequently self-limited, but immunocompromised persons may have more serious and prolonged infection (Benenson, 1995). Hospitalizations and deaths due to giardiasis are relatively rare; CDC estimates that giardiasis causes approximately 10-30 deaths and 3,500- 5,000 hospitalizations annually in the United States (Mead et al., 1999; Scallan et al., 2011a).

In some patients, symptoms last for only three or four days, while in others the symptoms last for months. Generally, patients commonly resolve their infections spontaneously, with acute disease lasting from one to four weeks (Smith and Wolfe, 1980). However, in some patients, the acute stage may persist for months (Wolfe, 1990). The period of communicability lasts for the entire duration of infection; however, the shedding of cysts can be intermittent (Benenson, 1995).

Clinical data suggest that *Giardia* cysts are highly infective for humans (Rendtorff 1954a,b, Rendtorff and Holt 1954a,b; Rose et al., 1991; Teunis et al., 1996).

Isospora belli

Transmission is through fecal contamination of water and food from a human source, especially in developing nations and in tropical regions (Marshall et al., 1997). Clinical manifestations of *I. belli* infection, isosporidiosis, are most commonly chronic to severe diarrhea, which can persist for months to years, causing weight loss, abdominal colic, and fever (Garcia, 2006b). These health outcomes are generally increased in severity and duration in immunocompromised persons. Worldwide, *I. belli* have been attributed to at least three documented waterborne disease outbreaks (Karanis et al., 2007)

Microsporidia

Microsporidiosis is an emerging and opportunistic infection and is associated with a wide range of clinical, often organ-specific syndromes in humans, including diarrhea, hepatitis,

keratoconjunctivitis, renal failure, and even blindness (Didier et al., 2004; Cali, 2006; Didier and Weiss, 2006). However, persistent or self-limiting diarrhea are the most common symptoms associated with microsporidiosis in immune-competent or immune-deficient individuals, respectively (Didier et al., 2004). Microsporidia have been attributed to at least one recent waterborne disease outbreak (Karanis et al., 2007).

Toxoplasma gondii

In humans, *T. gondii* is normally transmitted by ingestion of food or water contaminated with oocysts or by ingesting infective animal tissues that contain tissue cysts (Dubey, 2004, 2006). In pregnant infected hosts, *T. gondii* can multiply in the placenta and spread to fetal tissues, especially during the first half of gestation, causing mental retardation, loss of vision, hearing impairment, and mortality in congenitally infected children. Although toxoplasmosis is mostly an asymptomatic infection in adults, it can cause serious disease morbidity and mortality in immunocompromised persons (especially encephalitis in AIDS patients) (Artigas et al., 1994; Sanchez et al., 2000).

4.1.4. Extent of Secondary Transmission

Whether secondary transmission is included in the analysis needs to be determined during problem formulation. If the questions the risk assessment is asking requires secondary transmission to be considered, then dynamic MRA models can characterize secondary cases that occur among contacts following exposure to a primary case. Static MRA models usually consider secondary transmission to be negligible or include it as a non-fluctuating multiplicative factor (e.g., secondary cases equal primary cases multiplied by 0.1; assuming a 10% secondary transmission rate).

When secondary transmission is included the immune status of individuals becomes more important for modeling. For example, individuals who are already infected with a particular pathogen should not be considered susceptible to reinfection by the same pathogen while infected. Previously infected but recovered individuals have decreasing immunity over time, which may affect the design of the risk assessment model. Section 2.3.1.2 includes more information on dynamic models and the inclusion of secondary transmission in risk assessment.

4.2. Dose-Response Assessment Overview

In the case of waterborne microbial contaminants, risk assessment generally involves estimating the probability of illness or infection based on exposure estimates and dose-response relationships. During dose-response analysis, data from human clinical studies, epidemiological studies, animal studies, and/or outbreaks are used to develop a mathematical relationship between the intensity of exposure and the subsequent occurrence of disease or infection.

Dose-response models are generally derived as the logical mathematical consequence of the assumptions made about the infection process. For example, the exponential dose-response relationship is derived by assuming that the distribution of organisms between doses is random (i.e. Poisson), that each organism has an independent and identical survival probability, and that a single organism can cause infection. The form of the relationship between exposure and response is determined by (1) assumptions related to the biological processes leading to infection, and (2)

the "shape" of the relationship found in the data between exposure and the health outcome of interest. Statistical techniques are often used to optimize the relationship's parameters to best fit the available data.

A number of factors are addressed in the derivation of dose-response models and estimation of their parameters for microbial risk assessment, including the following (adapted from ILSI, 2000):

Dose-response factors

- statistical model(s) to analyze or quantify dose-response relationships;
- human and/or animal dose-response data;
- source and/or preparation of challenge material or inoculums.

Factors that overlap with exposure assessment

- utilization of outbreak or intervention data (can also be used to build exposure scenarios);
- route of exposure or administration used in the dose-response study;
- equivalence of methods used (including organism type, strain, and method units) for occurrence data and dose-response study;

Factors that overlap with health effects

- characteristics of the exposed population (age, immune status, etc.); and
- Infection or disease endpoint for the dose-response relationship (e.g., pathogen shedding, serological response, symptoms).

The mathematical form of the dose-response model may vary with pathogen or strain, route of administration, distribution of host statuses, and other factors. An overview of common dose-response models for microbial based infections is provided below. Either human or animal data may be used to derive dose-response estimates—although human data are generally preferred if they are available. Pooling data from different animal models can enhance dose-response models (Bartrand et al., 2008). Although pooling animal and human data should be conducted with caution (FDA/USDA, 2003). Information from disease outbreaks may also provide useful information about both primary infection risks (infections arising directly from exposure) and secondary transmission (person-to-person) (Teunis et al., 2004, 2008a).

Knowledge of the conditions under which dose-response data were collected is essential both for those developing dose-response models and for those evaluating dose-response models for use in MRA. In particular, the strain of the pathogen, model form, enumeration method, and route of inoculation can strongly influence the use of a reported dose-response relationship. Extrapolation of dose-response relationships to conditions other than those for which data were collected should be done only in conjunction with justification and with a full description of the conditions for which the dose-response model was developed.

For example, dose-response models have been proposed based on data collected during experiments with animal hosts whose response appears to differ substantially from that of humans.

Comparison of murine response to *Listeria monocytogenes* (based on feeding studies) to that of humans (based on epidemiological data) indicates a factor of ~10^6 difference in the lethal dose for 50 percent of the fetal population (LD_{50}) between the two hosts (FDA/USDA, 2003). Numerous, albeit dated, studies of pathogen-host combinations (e.g., Bell et al., 1955, for tularemia; Holdenfried and Quan, 1956, for plague) have shown the potential for wide variation in response between animal hosts and even among animal hosts of the same species but of different geographic origin. Insights into the applicability of animals as models of human infection may sometimes be drawn from pathology literature; in many cases animal models are selected for pathology experiments based on similarities between their infection process and that of humans (Lyons and Wu, 2007). Knowing and reporting these similarities may provide information for interpretation of dose-response data. As another example, if a QMRA were to use rotavirus data collected via a molecular method, such as reverse transcriptase (RT-)PCR methodology, it would be necessary to harmonize those data with the median tissue culture infective dose data reported during the clinical trial used as the basis for the commonly used dose-response relationship.

In summary, the use of dose-response models developed based on animal data for estimating human response is not universally accepted within the risk community. Reasons for using such models are that researchers select animal models based on similarity in response to that of humans, because uncertainties related to extrapolating animal models to humans may be less than uncertainties inherent in other techniques for developing human dose-response models (e.g., dose uncertainties when outbreak data are used to generate dose-response models), and because, in some cases, human data are not available at the low doses of concern and animal data can be obtained whereas human data cannot. Current research on dose-response models may provide an avenue for verifying the appropriateness of using animal dose-response models to estimate human response.

Different strains of a pathogen may have very different degrees of infectivity and the strain or strains used to generate the experimental data may or may not be the most relevant to the population that is exposed in the exposure scenario that is being modeled. For example, the methods used to determine doses for dose-response relationships may vary in their sensitivity for detection and hence introduce uncertainty into the accuracy of doses delivered (e.g., different cell lines used for infectivity-based assays for viral titres vary in their susceptibility to infection). Furthermore, and as noted previously, susceptibility in the general population may vary greatly with age, general immune status, and pre-existing or acquired immunity to specific pathogens and pathogen strains. The narrative accompanying the modeling should explain the potential impacts of different strains (or other sub-classifications of pathogens such as serovar or single nucleotide polymorphisms) on the outcome of the risk calculation. The potential uncertainty associated with variation in infectivity is rarely included in MRA studies currently; however, it should be calculated quantitatively if possible. Most MRAs treat severity of symptoms independently of dose size because most of the established dose-response relationships are based on infection rather than illness. However, there is some evidence for dose-dependence of severity of symptoms (Text Box 6).

Text Box 6. Dose-Dependency of Host-Pathogen Interactions

Most MRAs assume that severity of health endpoint is not influenced by magnitude of dose. For example, with *Cryptosporidium*, whether an individual is exposed to 1 or 10,000 organisms, if they become infected and ill, the health end points are assumed to be similar in severity. Thus, the assumption is made that exposure to a larger *Cryptosporidium* dose will not result in worse symptoms (U.S. EPA, 2006a). Although there is emerging evidence that this assumption may not be appropriate for all pathogens at all doses, the data are generally insufficient to be included in quantitative risk assessment. However, if there is evidence of dose-dependent severity of symptoms for the pathogen of interest, then it should be discussed. For example, pathogenic *E. coli* feeding studies in human volunteers demonstrated dose-dependency of disease severity and suggested that volume of liquid stool can be used as a quantitative metric for illness severity (Bieber et al., 1998). Colwell et al. (2003) reported that in Bangladeshi villages where sari or nylon cloth was used to filter surface water, the number of cholera cases was reduced and the severity of disease was reduced compared to villages that did not filter their surface water. Nauta et al., (2009) compared six *Campylobacter* risk assessments and concluded that the most effective public health intervention measures for risks associated with exposure to broiler chickens targeted *Campylobacter* density reductions, rather than reducing its prevalence (note that this finding may or may not extend to environmental waters).

The duration of exposure, the number of exposures, and time between exposures may affect the probability of an adverse health effect, as estimated by the dose-response relationship. As discussed previously, determining the independence or lack of independence of exposure events is complicated by host status as well as pathogen characteristics. For example, individuals who are already infected with a particular pathogen should not be considered susceptible to reinfection by the same pathogen while infected. Previously infected but recovered individuals have decreasing immunity over time. However, the nature of the decrease in immunity depends on many conditions including whether subsequent exposures boost immunity, host factors that relate to overall health of the immune system, and pathogen factors such as rapidly evolving antigenic epitopes.

Because any given person's immunity fluctuates based on many host factors, and because different pathogens elicit different immune responses, it is difficult to define a single exposure duration that best describes all combinations of host-pathogen interactions. Given the variability and complicated nature of capturing all appropriate exposure durations, most risk assessments choose a default exposure event duration (e.g., all water consumed during 1 day).

4.2.1. Overview of Common Dose-Response Model Forms for Pathogens

This section provides a brief summary of the most commonly used dose-response models for microbial pathogens. Namata et al. (2008) also provides a useful summary of dose-response models for MRA. Although providing state-of-the-art guidance on deriving dose-response relationships is beyond the scope of this MRA Tools document, there are several issues that risk analysts and risk managers should be aware of when evaluating the dose-response literature. Appendix C provides additional information on dose-response modeling.

The objective of the dose-response assessment is to develop a relationship between the number of microbes a person or population has been exposed to and the likelihood of occurrence of an adverse consequence (health outcome). In general, dose-response assessment would be relatively straightforward if the level of microbial risk that was deemed acceptable was sufficiently high to

allow experimentation that would permit the directly observable assessment of risk (Haas et al., 1999). However, the probability of infection (risk) from a single low-dose exposure event is often sufficiently low that use of direct observation (or experimentation) is impractical. For example, using a standard dose-response model (Medema et al., 1996), a dose of 30 *Campylobacter jejuni* corresponds to an infection probability of ~0.2, but that probability rises to 0.5 at a much higher dose (~900). Thus, the use of parametric dose-response curves to facilitate extrapolation into the low-dose range that matches the risk level of concern is necessary.

Dose-response models are mathematical functions that input the dose to which individuals or populations are exposed and yield a probability (bounded by 0 and 1) of the particular adverse health effect (Haas et al., 1999). These dose-response functions play a prominent role in risk assessments for pathogens in water because they effectively translate exposures into risks. In real world situations where large numbers of individuals may be exposed (e.g., public water supplies), relatively low individual risk levels may be of concern from a public health perspective because even low individual risks can translate into a large number of illnesses.

The two most commonly used dose-response models are the exponential and beta-Poisson models. The use of exponential and beta-Poisson models is only valid, however, when their underlying assumptions are met. More computationally intensive dose-response relations are also available for conditions in which neither the exponential or beta-Poisson models are appropriate. Alternative two-parameter models have been proposed for use in MRA assessment, including the log-normal, log-logistic, extreme value models (Pinsky, 2000). Three-parameter models that have been suggested for MRA include the Weibull gamma (Farber et al., 1996), exponential gamma, Weibull exponential, and the shifted Weibull model (Kodell et al., 2002). Although three-parameter models are more flexible than two-parameter models, they require data at four or more doses, which is not available for many microbial pathogens. Research continues to be conducted on appropriate methods for selection of models from among these and other candidate models (e.g., Moon et al., 2004, 2005).

The models discussed in this section estimate risks for exposed individuals. Population-level risks (i.e., the incidence of disease among a group of exposed individuals) are generally constructed by combining individual risks with estimates of the distribution of doses to the exposed population.

To promote transparency and clarity in an MRA, the following points should be addressed for each dose-response model chosen:

- a discussion of assumptions inherent in making extrapolations to doses lower than those used in studies;
- a detailed description of dose-response and risk assessment modeling approaches, including the applicability of the models for use in various exposure situations and for various pathogens;
- methods used to assay doses and exposure, because they may not be equivalent, and a summary of approaches taken to harmonize between methods;
- the models' key assumptions;
- the type of information that the various models are expected to provide;
- limitations of the models;

- the use of likelihood methods to compare how well dose-response models fit the data;
- the biological rationale for the model selected;
- the strengths/weaknesses and advantages/disadvantages of the models, including a comparison of the benefits and limitations of the chosen models versus other potential models; and
- a discussion of the flexibility in approaches to the dose-response relationship depending on the pathogen being considered and the assumption about a no-threshold effect (i.e., can it be assumed that one organism is sufficient to produce infection in some portion of an exposed population or subgroup?).

To take biological mechanisms into account, a dose-response model for microbes should account for the heterogeneous distribution (random or clumping) of microbes in water (affecting exposure) and a microbe's ability to reproduce in the human body (linked to pathogenicity) (Haas et al., 1999). Laboratory dose-response studies are usually conducted under conditions in which the microorganisms are randomly distributed in the administered dose. This is known as a Poisson distribution. The framework for exponential models is based on well-studied mathematical relationships; however, the model parameters use empirical data from clinical trials and epidemiological studies that are organism-specific (e.g., an organism's infectious dose for 50% of the exposed population [ID_{50}]). A concern for environmental water samples regarding the Poisson distribution is clumping, association with suspended solids, and other spatial distribution issues; however, this phenomenon can be accounted for in dose-response modeling by incorporating aggregation parameters into the dose- response model (see Teunis and Havelaar, 2000 and Teunis et al., 2008b for further information).

The dose-response relationship that is defined by the equation is "fit" to experimental data using a variety of statistical methods. If the model is a good fit, it will predict risks that are close to those actually observed within the range of experimentally administered doses. However, the doses used in volunteer studies may be higher than those typically encountered in the environment, so it is necessary to extrapolate the risks associated with lower doses using the model derived from the higher doses. In extrapolating to lower doses, risk assessors rely on the belief that the form of the dose-response model is based on an accurate representation of the infection process that holds at low doses as well as high doses. Text Boxes 7 and 8 illustrate how experimental data on *Cryptosporidium* and noroviruses have been used to derive dose-response relationships for these pathogens. Moreover, these dose-response relationships are generally based on clinical trials for which only the average doses are known for each group. Recently, outbreak data have been used to derive dose-response relationships for several waterborne pathogens. Teunis et al. (2004, 2008a, 2010) and Bollaerts et al. (2008) provide good examples of how outbreak data have been used to derive dose-response relationships.

Several published studies (e.g., Coleman and Marks, 2000; Nauta et al., 2009) suggest that it might not be advisable to extrapolate dose-response models based on clinical trials for waterborne exposures, given the complexity of the pathology of illnesses and given the relatively low reported incidence of illness and the relatively high daily exposure of humans to pathogens (Levin and Antia, 2001). Although a critical evaluation of this perspective is difficult to provide, given the limited data available for human response to exposure to pathogens of known dose and

characteristics, mechanistic modeling offers an avenue for development of improved models for extrapolation to low doses.

Text Box 7. Brief Summary of *Cryptosporidium* Feeding Studies

Human feeding studies have been used for decades to systematically evaluate dose-response effects for pathogens. Chappell et al. (Chappell et al., 1999, 2006; DuPont et al., 1995; Okhuysen et al., 1999, 2002) have conducted volunteer feeding studies using five strains of Cryptosporidium parvum that have formed the basis of dose-response parameters used in several MRAs. Students and employees of the University of Texas Health Science Center and others in the surrounding area of Houston were recruited as volunteers. Volunteers could not be caretakers of infants, elderly, or those with chronic diseases or immunosuppression. Recruits were given extensive information on cryptosporidiosis and had to score 100% on a written examination that tested the recruits on their comprehension of the study, the fact that they could become ill, that there was no effective treatment for the illness, and that the organisms could be spread to household contacts. The next stage of their evaluation for inclusion in the study involved providing medical histories and passing extensive medical tests.

In each of the three studies, the volunteers ingested a single known dose of viable *C. parvum* oocysts of one of three isolates—IOWA, TAMU, and UCP. The subjects were given anywhere from 10 to 1,000,000 oocysts per dose.

Volunteers submitted stools passed after the challenge and completed daily diaries regarding their stool passage and any symptoms; household contacts were also monitored for diarrheal illness. Blood was collected from each of the volunteers at specified days post-challenge and tested for antibody response.

To measure the challenge responses, the researcher considered two definitions of infection—confirmed and presumed. A confirmed infection was based on oocysts detected in stools using direct fluorescence assay. Some volunteers who had oocysts in their stools did not develop any symptoms. In contrast, some volunteers had illness symptoms that were indistinguishable from those with confirmed infection, but had no detectable oocysts in their stools. Because of the detection limit of the assay methodology, these volunteers were presumed to be infected. Those volunteers who did not have either GI symptoms or fecal oocysts throughout the study were presumed to be uninfected.

Because the purpose of the studies was to develop dose-response curves for the different strains based on infectious dose, it was important to be able to capture data on the median infectious dose. The first study of the IOWA isolate was designed to cover a wide range of doses (30-1,000,000 oocysts) so to more effectively capture the median dose. The doses of the other two isolates were adapted as the study progressed to narrow the range of doses; that is, the first group of volunteers were challenged at a moderate dose, while the next group's dose level was altered, depending on the outcome of the previous group. That way, the median infectious doses could be captured over a smaller dose range. The entire time required for each dose-response study was 11 to 14 months.

The infectivity for each of the three isolates was estimated using the study data and then tested using the exponential model (Messner et al., 2001). A comprehensive dose-response evaluation was conducted by EPA during the development of the LT2 (U.S. EPA, 2003a,b, 2006a). Teunis (2009) analyzed all five strains and found most likely ID_{50} appears to be in the range of 30 to 50 oocysts.

Text Box 8. Brief Summary of Challenge Studies to Investigate the Dose-Response and Host-Immunity Factors Related to Norovirus Infection

The group of viruses called norovirus (previously known as Norwalk or Norwalk-like viruses) is the most common cause of AGI outbreaks (e.g., through ingestion of contaminated food and water) in the United States (cite Scallan et al., 2011a,b). Reports implicated noroviruses in 94% of U.S. nonbacterial gastroenteritis outbreaks from 1996 to 1997 (Fankhauser et al., 1998). However, host immunity to noroviruses remains poorly characterized. Although over 70% of U.S. adults have serum antibodies to norovirus, the antibodies do not appear to confer any protection from reinfection (Greenberg et al., 1979). Despite outbreak studies suggesting that norovirus has high infectivity and high person-to-person transmissibility, certain exposed people never develop illness (i.e., remain asymptomatic).

Lindesmith et al. (2003, 2005) conducted a series of human volunteer studies to examine the dose-response characteristics of different strains of noroviruses (Norwalk [NV] and Snow Mountain Agent Virus [SMV]) and the role of host immunity in the probability of (re)infection. For these studies, the researchers recruited healthy adult volunteers with and without pre-existing serum immunoglobin G (IgG) to NV. The first study included 31 volunteers and examined 3 low doses of NV. The second study included 15 volunteers and examined 3 doses of SMV. The inoculum was diluted in sterile water and ingested. The volunteers stayed 5 consecutive days/6 consecutive nights at a research center for monitoring of GI symptoms, then reported for follow-up visits for collection of stool, serum, and saliva samples on days 8, 14, and 21 post-challenge.

Infection was defined as detection of viral shedding in stool by reverse transcriptase PCR (RT-PCR) or seroconversion designated by a 4-fold or more rise in the specific IgG. Symptoms that defined GI illness were diarrhea (defined as more than 2 unformed stools within 24 hours), vomiting, abdominal pain, muscle pain, fatigue, and chills—fever and headache were excluded.

Previous research has reported an association of a mutation in the alpha (1,2) fucosyltransferase gene (FUT2) gene with immunity to NV infection (Marionneau et al., 2002). In that study, volunteers with the FUT2 mutation remained healthy and had no significant increase in anti-NV salivary antibody titers, even after high-dose exposure. Note that about 20% of the North American population has the FUT2 mutation. Of the volunteers with fully functioning FUT2 genes, about half became infected. In the remaining uninfected half of the group, salivary IgA levels showed mucosal immune response post-challenge, suggesting that previous exposure had resulted in protective immunity.

Moe et al. (2002) found that infected subjects were generally older than uninfected subjects, and were twice as likely to have NV-specific IgG in their baseline serum specimen. Consequently, the presence of anti-NV serum IgG was not protective against infection. In other words, although these individuals had been exposed previously to NV, perhaps multiple times, they continued to be susceptible to reinfection. These studies provide important implications for microbial risk assessors—even with a very low infectious dose in susceptible populations, susceptibility to Norwalk virus is multifactorial and influenced by both acquired immunity and genetic traits.

In subsequent work, these researchers along with Teunis et al. (2008b) developed a dose-response relationship for NV based on challenge study data and a new variant on the hit theory model of microbial infection. This relationship accounts for variation in NV infectivity, as well as the degree of virus aggregation. Moreover, the results indicate that passage through a human host does not change NV infectivity and that NV is a highly infectious microorganism.

Exponential Model

The exponential model is the simplest model that is commonly used in MRA (Text Box 9); it is based on the following assumptions (Haas et al., 1999):

- microorganisms are distributed in water randomly[21] and thus, follow the Poisson distribution;
- for infection to occur, at least one pathogen entity must survive within the host; and
- the probability of infection (in a person or animal model) per ingested or inhaled organism is constant.[22]

Text Box 9. Summary of Use of Exponential Model (Source: Rose et al., 1991)

Rose et al. (1991) used an exponential model to estimate the risk of infection after exposure to treated water contaminated with Giardia cysts as shown by the following equation:

$P_i = 1-\exp(-r\mu V)$,

where P_i is the probability of infection, r is the host-pathogen interaction probability, μ is the average number of organisms, and V is the volume of water consumed.

The parameter designating the infectivity of *Giardia* (r) in the exponential model was based on data from studies in which volunteers were fed a range of 1 to 10^6 cysts and the response was measured by the number of cysts excreted in the volunteer feces, not by clinical symptoms (Rendtorff, 1954a,b; Rendtorff and Holt, 1954a,b). An average r value (the fraction of microorganisms that are ingested that survive to initiate infection) was compared by determining the value of r at each dose. Based on the results, the average r was calculated to be 0.01982.

Using the exponential model, the potential risk of infection was determined with varying levels of *Giardia* cysts in drinking water. The model used 2 L of unboiled water a day as the consumption parameter, V. The number of cysts (μ) was based on densities measured in source waters with 99.9, 99.99, and 99.999% estimated removal by treatment. The exposure was based on the numbers of cysts per L multiplied by 2L. A maximum daily risk was estimated using the highest level of contamination and a yearly risk was based on 365 days of exposure to the geometric mean density of cysts. The model was checked for plausibility by entering the data from five waterborne giardiasis outbreaks using the levels of *Giardia* cysts and the observed attack rates in the exposed population.

Under the exponential model, there is no minimum infectious dose, as a nonzero risk is predicted with any non-zero dose. Assuming that a single organism is sufficient to cause infection, and that the ingested organisms must pass through "multiple barriers" to survive long enough to cause disease, yields the exponential risk model:

$$P_r = 1 - e^{-rD}, \qquad [4\text{-}1]$$

[21] As noted previously, because microbes are generally not thought to be distributed randomly in environmental media, this assumption is considered to be a limitation of the exponential model unless adjustments are made as discussed in Teunis and Havelaar (2000).

[22] This assumption also introduces uncertainty because host variation is not considered.

Where:

- P_r is the probability of an individual exhibiting the response specified by dose-response parameter given exposure to the defined dose (unitless);
- D is the mean dose of microbes for a group in the clinical trial (number of microbes ingested per event, which is often represented as daily water intake × density);
- *r* is the dose-response parameter that is "fit" to the data; higher value indicates higher risks at lower doses[23]; and
- *e* is the base of the natural logarithm function (unitless).[24]

The "response" to exposure may be the development of clinical symptoms, and/or microbiological or immunological evidence that microbes have persisted or multiplied in the body. For example, oocyst shedding (regardless of whether illness symptoms are present) in stool is commonly used to indicate *Cryptosporidium* infection.

Under this model the median infectious dose is $N_{50} = \ln(2)/r = 0.693/r$ (where l is the natural logarithm).

Beta-Poisson Model

The beta-Poisson model is based on similar assumptions to the exponential model except that the third assumption (that the probability of infection per ingested organism is constant) is relaxed. This model allows the probability of infection per ingested or inhaled organism to vary within the exposed population (Haas et al., 1999). In this model the probability of surviving and reaching a host site (*r* in the exponential model) is beta distributed, and thus the model contains the two parameters (α and β) of the beta distribution. Thus, the beta-Poisson accounts for differential immunity in a population (but not specifically for differences between groups or subgroups in a population). The exponential model generally provides a good fit to experimental data if the infectivity of the administered organisms and the inherent susceptibility of the exposed population (animal or human) are constant. However, when there is variability in the host-pathogen interaction, diversity in the pathogen (as when multiple strains are present), or both, the dose-response relation tends to be shallower than that of the exponential relation. The most commonly used approximation to the beta-Poisson model is as follows:

$$P_r = 1-(1+D/\beta)^{-\alpha} \qquad [4\text{-}2]$$

Where:

- P_r is the probability of an individual exhibiting the response specified by dose-response parameter given exposure to the defined dose (unitless);
- D is the mean dose of microbes ingested (number of microbes ingested per event, which is often represented as daily water intake times density);

[23] For small values of r, the estimated individual risk is r×d (i.e., the model is linear at low doses).
[24] Euler's number or Napier's constant is ~2.71828.

- β is the location parameter; determines inflection point of dose-response curve (unitless); and
- α is the shape parameter governing the steepness of the dose-response curve (unitless).

Unfortunately, in this approximation to the beta-Poisson model, α does not have an obvious physical interpretation. What can be said is that it is a shape parameter governing the steepness of the dose-response curve; the larger its value, the steeper the curve (McBride et al., 2002). The derivation of the approximation to the beta-Poisson model, as shown above requires that β>>1, β >> α, and becomes a poorer approximation at small values of β or large values of D. In practice, this condition is not always met, and caution is warranted, especially for uncertainty calculations at low doses when this approximation is used (Teunis and Havelaar, 2000). This approximation to the beta-Poisson is linear at low doses and the curve is always shallower than the exponential model. However, as α approaches ∞, the approximate beta-Poisson model approaches the exponential model (Haas et al., 1999).

Under this model the median infectious dose is $N_{50}= \beta \times (2^{1/\alpha} -1)$.

When possible, it is preferable to directly fit the exact model (Equation 4-3), where ${}_1F_1(\alpha, \alpha + \beta, -D)$ denotes a confluent hypergeometric distribution with the specified parameters.

$$P_r = 1 - {}_1F_1(\alpha, \alpha + \beta, -D) \qquad [4\text{-}3]$$

Where:

- P_r is the probability of an individual exhibiting the response specified by dose-response parameter given exposure to the defined dose (unitless);
- D is the mean dose of microbes ingested (number of microbes ingested per event, which is often represented as daily water intake times density); and
- ${}_1F_1(\alpha, \alpha + \beta, -D)$ is Kummer's confluent hypergeometric distribution, with parameters α, α + β, and -D, variable means are analogous to the beta-Poisson approximation.

Although the hypergeometric distribution generally has no analytical solution, numerical estimation algorithms have been developed.

Bayesian Methods

Bayesian methods to estimate dose-response model parameters are also being increasingly used (Messner et al., 2001; Englehardt, 2004; Englehardt and Swartout, 2006). In general, a dose-response function gives the probability of illness or infection as a function of the dose and of several unknown parameters. Experimental data are collected from subjects accidentally or deliberately exposed to a measured microbial dose. The numbers of subjects that become infected or ill for each dose level are observed, leading to a binomial likelihood, which is the probability of the observed numbers of cases given the unknown parameters. The "traditional" frequentist statistical approach uses the binomial likelihood, and chooses parameter values to maximize the likelihood. Uncertainty intervals for the parameters are called confidence intervals (i.e., on average, out of 100 95% confidence intervals, 95 will contain the parameter value) (see Section

5.3.1). Confidence intervals around the maximum likelihood estimates can be calculated using approximations valid for large samples. For small sample sizes that are typical in MRA, bootstrap confidence intervals are calculated by randomly resampling from the original dose-response data and estimating the parameters for each of these bootstrap samples. However, if the sample sizes are too small, bootstrapping becomes less desirable (Teunis and Havelaar, 2000).

Bayesian methods, which are discussed in more detail in Appendix C, exploit available subjective and related information in addition to the numeric data. Ideally, the investigator expresses an initial assessment of the unknown parameter distribution before examining the data by defining a prior probability distribution for the parameters. The prior probability distribution is defined based on subjective information and professional judgment. Recently published MRAs have used a "non-informative" prior distribution to represent the lack of prior information. Using Bayes' rule, the posterior probability distribution for the parameters given the data can be calculated. In a Bayesian analysis, uncertainty intervals for the parameters and the dose-response function can be calculated from the posterior distribution as "credible intervals"; for example, a 95% credible interval has a 95% probability of including the parameter value, given the data.

The MCMC method is often used to simulate values from a posterior probability distribution for which direct analytical calculations are difficult, intractable, or inconvenient. Gilks et al. (1996) provides a good description of these methods. Instead of being statistically independent, the consecutive values form a Markov Chain so that the statistical distribution for one value depends upon the previous value. After a sufficiently long "burn-in" period, every k^{th} value is sampled, giving an approximately random sample from the posterior distribution. It is unnecessary to know the normalizing constant that makes the distribution integrate to one. A version of the Metropolis-Hastings algorithm (Hastings, 1970; Gilks and Wild, 1992; Gilks et al., 1996) is used at each step to simulate from the posterior distribution without knowing the normalizing constant.

An advantage of the Bayesian approach over the frequentist approach is the ability to incorporate prior information. However, for the MRAs in the current literature this is not very useful because the prior information is too limited and non-informative priors have been used. The subjective nature of the choice of prior distribution is often thought to be a disadvantage of the Bayesian approach. A more important advantage of the Bayesian approach is that, unlike frequentist confidence intervals, the uncertainty intervals from a Bayesian analysis are easier to interpret and are usually interpreted correctly. Furthermore, the Bayesian uncertainty estimates of dose-response functions are generally easier to calculate and more exact than the frequentist confidence intervals. Finally, Bayesian methods are well-suited to meta-analysis of multiple studies, pathogens, or populations.

A predictive Bayesian dose-response function can be developed as follows. First, the parametric form of the dose-response function is established by theoretical derivation and, if possible, empirical confirmation. Then all available knowledge, other than the theoretical form of the conditional distribution and empirical data already used for that purpose, is considered during estimation of the parameters of the distribution. To do this, the parameters are recognized as uncertain but subject to professional judgment, and thus, a prior probability distribution is assigned to each parameter. Prior distributions are then refined with dose-response data, to obtain a posterior distribution. Next, the predictive Bayesian dose-response function can be found by multiplying the

posterior by the conditional dose-response function and integrating over the parameter space (Englehardt, 2004). As noted previously, MCMC methods can then be used to generate samples from the joint posterior distribution (Messner et al., 2001).

Several researchers advocate for the combined use of Bayesian and frequentist (likelihood-based) methods (Teunis and Havelaar, 2000; Messner et al., 2001). Often the frequentist approach is used to provide maximum likelihood estimates of the dose-response function, and the Bayesian approach is used to calculate uncertainty intervals (e.g., 80% or 95% credible intervals for the parameters or the dose-response). Several papers use the Bayesian posterior mode to select the dose-response function (Teunis et al., 2004, 2005, 2008a,b). The posterior mode is given by the parameters that maximize the posterior probability, defined as the product of the prior and the likelihood; thus, it is not necessary to calculate the normalizing constant for this calculation.

Other Dose-Response Methods

In addition to the dose-response models described above (exponential, beta Poisson), there are other dose-response models either in use in QMRAs or that can be potentially incorporated into QMRAs. These models include empirical dose-response models, threshold models, and mechanistic models of varying resolution. The exponential and beta-Poisson models are distinguished from empirical models because their derivation is based on a sequence of plausible events, although this assessment is not universal (e.g., see Coleman and Marks, 1998). Threshold models have in some cases provided significant improvements in fit over the exponential and beta-Poisson models, but their use has been advocated on the basis of analysis of the infection process and interpretation of epidemiological data. Mechanistic models are currently in development and offer the potential for development of dose-response models for pathogens for which dose-response data are unavailable of for the low-dose range. These models depict the pathogen-host system in varying resolutions and may be stochastic, deterministic, or a combination. Given the widespread use of the exponential and beta-Poisson dose-response models for waterborne pathogens and the advantages these models offer (as described above), these alternative empirical, threshold, and mechanistic models are not presented in the body of this report, but are described and contrasted with the exponential and beta-Poisson model in Appendix C.

4.2.2. Summary of Available Dose-Response Relationships for Waterborne Pathogens

An overview of the many of the currently available and generally accepted dose-response relationships for waterborne pathogens is summarized in Table 10, which includes the pathogens listed alphabetically, the resulting dose-response form and parameter values, and the corresponding reference for that work. Note that in some cases reported relationships are also provided for illness conditional on infection. Use of these conditional relationships is typically not universal in MRA modeling, as often illness is modeled as variable independent of dose.

Table 10. Overview of Dose-Response Relationships and Health Effects for Waterborne Pathogensa (Source: Adapted from McBride et al., 2002)

Microorganism	Model	Parameters[b]	Health Effects	Reference(s)
Adenovirus 4	Exponential	$r = 0.4172$[c]	Acute infectious nonbacterial GI illness, acute febrile respiratory disease, acute hemorrhagic conjunctivitis, phary-noconjuctival fever	Crabtree et al., 1997 Haas et al., 1999 APHA, 2004
Campylobacter jejuni[h,i]	Beta-Poisson	$\alpha = 0.145$ $\beta = 7.59$	Diarrhea, abdominal pain, malaise, fever, nausea, and vomiting (typhoid-like syndrome), febrile convulsions, meningeal arthritis, reactive arthritis, Guillain-Barré syndrome)	Medema et al., 1996 Teunis et al., 1996 Haas et al., 1999
	Infection: Hypergeometric beta-Poisson (for healthy adults)	$\alpha = 0.024$ $\beta = 0.011$		APHA, 2004 Teunis et al., 2005
	Illness: Conditional on infection (for children)[g]	$R = 2.44 \times 10^{8}$ $\eta = 3.63 \times 10^{-9}$		
Coxsackievirus	Exponential	$r = 0.0145$	Vesicular pharyngitis (acute self-limited, viral disease characterized by sudden onset, fever, sore throat and small pharyngeal lesions)	APHA, 2004 Haas et al., 1999
Cryptosporidium[d]	Exponential	$r = 0.0042$ Iowa isolate	Cryptosporidiosis: profuse watery diarrhea, malaise, fever, anorexia, nausea, and vomiting	APHA, 2004 Haas et al., 1996, 1999
		$r = 0.077$		Okhuysen et al., 1999
		$r = 0.0572$ TAMU isolate		http://wiki.camra.msu.edu
		r = in the range 0.04 to 0.16 for unknown mixture environmental isolates		U.S. EPA, 2006a, Dupont et al. 1995, Okhuysen et al., 1999, 2002, Chappell et al., 2006
	Beta-Poisson	$\alpha = 0.27$ $\beta = 1.40$ TU502 isolate $\alpha = 0.114$ $\beta = 1.04$ Moredun isolate $\alpha = 0.145$ $\beta = 1.52$ UCP isolate		http://wiki.camra.msu.edu
	Generalized beta-Poisson for Illness	$\alpha = 0.060$ $\beta = 0.095$		Englehardt and Swartout, 2006

Table 10. Overview of Dose-Response Relationships and Health Effects for Waterborne Pathogensa (Source: Adapted from McBride et al., 2002) (continued)

Microorganism	Model	Parameters[b]	Health Effects	Reference(s)
Echovirus 12[j]	Exponential	r = 0.0128	Acute Febrile respiratory disease	Haas et al., 1999 APHA, 2004
	Beta-Poisson	α = 0.401 β = 227.2		Teunis et al., 1996
		α = 0.374 β = 186.7		Regli et al., 1991 Rose and Sobsey, 1993
		α = 1.3 β = 75		Rose and Gerba, 1991
Endamoeba coli	Beta-Poisson	α = 0.1008 β = 0.3522	Not a human pathogen	Haas et al., 1999
Escherichia coli (pathogenic strains)	Beta-Poisson	α = 0.1778 β = 1.78×10^{6}	Acute watery diarrhea	Haas et al., 1999 APHA, 2004
E. coli O157:H7	Beta-Poisson[e]	α = 0.248 β = 48.80	Diarrhea (bloody), severe abdominal cramping, headache, hemorrhagic colitis, and hemolytic uremic syndrome	APHA, 2004 Teunis et al., 2008a
	Hypergeometric beta-Poisson	α = 0.084 β = 1.44 (children) α = 0.050 β = 1.001 (adults)		Teunis et al., 2004
Giardia lamblia	Exponential	r = 0.0199	Giardiasis: diarrhea (chronic); abdominal cramps; bloating, frequent loose, pale, greasy stools; fatigue; malabsorption	Regli et al., 1991 Rose and Gerba, 1991 Rose et al., 1991 Teunis et al., 1996 Haas et al., 1999 APHA 2004
Hepatitis A virus	Exponential	r = 0.5486[f]	Hepatitis: acute inflammation of the liver	Haas et al., 1999
Legionella	Exponential	r = 0.06	Legionellosis, pneumonia, Legionnaire's disease, Pontiac fever	Armstrong and Haas, 2008
Norovirus	Infection: Hypergeometric function ${}_1F_1$ (note: if aggregation is different than reported then ${}_2F_1$ function is needed) Illness: Conditional on Infection[g]	α = 0.040 β = 0.055 $\eta = 2.55 \times 10^{-3}$ r= 0.086	Usually self-limited, mild to moderate disease with clinical symptoms of nausea, vomiting, diarrhea, abdominal pain, myalgia, headache, malaise, low grade fever, or a combination of these symptoms	APHA 2004 Teunis et al., 2008b Messner et al., 2014

Table 10. Overview of Dose-Response Relationships and Health Effects for Waterborne Pathogensa (Source: Adapted from McBride et al., 2002) (continued)

Microorganism	Model	Parameters[b]	Health Effects	Reference(s)
Poliovirus III	Beta-Poisson	$\alpha = 0.409$ $\beta = 0.788$	Non-specific fever, acute flaccid paralysis	Rose and Sobsey, 1993 APHA, 2004
		$\alpha = 0.409$ $\beta = 0.788$		Regli et al., 1991
		$\alpha = 0.5$ $\beta = 1.14$		Rose and Gerba, 1991
Rotavirus	Beta-Poisson	$\alpha = 0.26$ $\beta = 0.42$	Sporadic, seasonal, often sever gastroenteritis of infants and young children, characterized by vomiting, fever and watery diarrhea	Gerba et al., 1996b, Haas et al., 1999 Regli et al., 1991 Rose and Sobsey, 1993
		$\alpha = 0.232$ $\beta = 0.247$		Rose and Gerba, 1991 APHA, 2004
	Hypergeometric beta-Poisson	$\alpha = 0.167$ $\beta = 0.191$		Teunis and Havelaar, 2000
Salmonella spp.	Beta-Poisson	$\alpha = 0.33$ $\beta = 139.9$	Gastroenteritis (enteric fever and septicemia)	Rose and Gerba, 1991 APHA, 2004
	Gompertz log (illness)	ln(a) in the range 29 to 50 b = 2.148		Coleman and Marks, 2000 Coleman et al., 2004 Soller et al., 2007
	Generalized linear mixed models and fractional polynomials of dose[k]	$\beta_0 = 0.323$ $\beta_1 = 5.616$ $\beta_2 = -8.462$ $\beta_3 = -7.782$ $d^2 = 0.780$		Bollaerts et al., 2008
Salmonella (non-typhoid)	Beta-Poisson	$\alpha = 0.3126$ $\beta = 2885$		Haas et al., 1999
	Bayesian mixed model : Infection: Hypergeometric function ${}_2F_1$ Illness: Conditional on Infection	$\alpha = 0.00853$ $\beta = 3.14$ (note 5000 samples of model parameters available from cited author) $\eta = 6.9\times10^{1}$ r= 8.23		Teunis et al. 2010
Salmonella typhi	Fractional polynomials	$\beta_1 = -18.1425$ $\beta_2 = 22.5300\times10^{-5}$		Namata et al., 2008
	Beta-Poisson	$\alpha = 0.1086$ $\beta = 6,097$		Haas et al., 1999
		$\alpha = 0.21$ $\beta = 5,531$		Rose and Gerba, 1991

Table 10. Overview of Dose-Response Relationships and Health Effects for Waterborne Pathogensa (Source: Adapted from McBride et al., 2002) (continued)

Microorganism	Model	Parameters[b]	Health Effects	Reference(s)
Shigella	Beta-Poisson	α = 0.21 β = 42.86	Shigellosis: acute gastroenteritis, dysentery, fever, nausea, vomiting, and cramps	Haas et al., 1999
Vibrio cholera	Beta-Poisson	α = 0.25 β = 16.2	Profuse, watery diarrhea; vomiting	Haas et al., 1999, APHA 2004

[a] Calibrations based on available data that have used particular pathogen strains processed in particular ways. Where more than one strain of an organism has been studied in clinical trials, a wide range of infectivities can be discovered. Therefore it must be recognized that these calibrations can carry a substantial degree of uncertainty.
[b] For the exponential distribution $N_{50}= 0.693/r$; for the beta-Poisson distribution $N_{50}= \beta * (2^{1/\alpha} -1)$.
[c] Developed for inhalation exposure to adenovirus 4 aerosols.
[d] Okhuysen et al. (1999) Estimated based on ID_{50} reported for the TAMU isolate. Teunis 2009 graphs 5 clinical trials resulting in an estimated ID_{50} of 30-50.
[e] Represents a meta-analysis of seven outbreaks and adjusting for heterogeneity. Alpha/beta pairs derived via MCMC analyses are available from Dr. Teunis. Use of those pairs is preferred to the use of the values shown in this table
[f] Corresponding dose units are grams of feces.
[g] Dose-response relation for the conditional probability of illness in infected subjects = $1 – (1+ \eta CV)^{-r}$, where η and r are shown in the table; CV is the dose (concentration × volume).
[h] An alternate dose-response model is proposed by Brynestad et al. (2008). That model is not included in Table 10; however, it is described along with other empirical models in Appendix C.
[i] Coleman and Marks (2004) suggest the dose-response models for *Campylobacter* identified in this table do not account for strain variability sufficiently and suggest the need for development of more detailed mechanistic models.
[j] Note that the dose-response models for Echovirus 12 yield very different ID_{50} values of ~54 or ~1053 depending on whether or not a factor of 33 is applied to the dose range. Teunis et al. (1996) fitted a beta-Poisson model to the echovirus 12 clinical trial data (Schiff et al., 1984) for 149 volunteers given doses of 0, 330, 1000, 3300, 10,000, 33,000, and 330,000 PFU. This model gives $ID_{50} \approx 1052$. Haas (1983) fitted a simple exponential model for the same virus to a set of clinical trial data reported by Akin (1981), in which 60 volunteers were given doses of 10, 30, and 100 PFU. That model, gives $ID_{50} \approx 54$ (Haas et al., 1999). The Akin data appears to be a preliminary subset of the Schiff data and there was only one clinical trial.
[k] Derived based on a series of foodborne outbreaks; not necessarily valid for water matrices or exposures.

4.3. Host-Pathogen Profile and Linkage between Human Health Effects Assessment and Other MRA Components

The host-pathogen profile is a distillation of the most important information and analyses that are conducted during the human health effects assessment. The host-pathogen profile can provide, depending on the available data, a qualitative and/or quantitative description of the human health effects scenario (ILSI, 2000). An assessment of the assumptions made during the human health effects assessment, and the uncertainty associated with the assessment because of lack of knowledge about the scenario or insufficient experimental or epidemiological data, should be presented. Any assumptions based on scientific judgment should be described and justified in the host-pathogen profile. A summary of the quantitative or qualitative uncertainty analysis should also be included.

Thus, the host-pathogen profile serves as the critical linkage from the human health effects assessment to the exposure assessment. The iterative nature of risk assessment requires that the host pathogen profile and the exposure profile be critically evaluated by the risk assessors and

managers to determine if the problem formulation component of the risk assessment needs to be revisited and refined based on the availability of relevant data presented in these profiles.

5. Risk Characterization

As noted throughout the preceding chapters, risk assessment is an iterative process. During risk characterization, the results of this iterative risk assessment process are integrated and documented. Thus, risk characterization is the culmination of the MRA process, and the final integrative step. The risk characterization needs to be complete, informative, and useful for decision-makers. The agency's *Science Policy Council Handbook: Risk Characterization* (U.S. EPA, 2000b) describes risk characterization as the step that "integrates information from the preceding components of the risk assessment and synthesizes an overall conclusion about the risk that is complete, informative, and useful for decision-makers" (U.S. EPA, 2000b, 2012a).

Risk characterization forms the starting point for formulating risk management considerations and provides a foundation for (regulatory) decision-making. It characterizes both quantitative and qualitative data in technical and non-technical terms, explaining the extent and weight-of-evidence, results, and major points of interpretation and rationale. It also summarizes the strengths and weaknesses of the evidence, conclusions, uncertainties, variability, potential impact of alternative assumptions, and discusses scenario, model, parameter, and analysis options that may deserve further consideration as the results from the assessment are subsequently used for decision-making purposes.

EPA's policy statement on risk characterization (U.S. EPA, 2000b) is as follows:

> Each risk assessment prepared in support of decision-making at EPA should include a risk characterization that follows the principles and reflects the values outlined in this policy. A risk characterization should be prepared in a manner that is clear, transparent, reasonable and consistent with other risk characterizations of similar scope prepared across programs in the Agency. Further, discussion of risk in all EPA reports, presentations, decision packages, and other documents should be substantively consistent with the risk characterization. The nature of the risk characterization will depend upon the information available, the regulatory application of the risk information, and the resources (including time) available. In all cases, however, the assessment should identify and discuss all the major issues associated with determining the nature and extent of the risk and provide commentary on any constraints limiting fuller exposition.

5.1. Introduction to Risk Characterization

The Agency's risk characterization policy (U.S. EPA, 2000b) calls for a transparent process and documentation that is clear, consistent, and reasonable. TCCR is particularly relevant for risk characterization because a risk assessment is often judged by the extent to which the risk characterization achieves the principles of TCCR. This section provides a summary overview that is intended to provide risk assessors, risk managers, and other decision-makers an introduction to the goals and principles of risk characterization. More comprehensive documentation on the topic of risk characterization has been prepared by the Agency and interested readers are referred to that documentation (U.S. EPA 2000b). This document complements and extends that previous work by discussing tools, methods, and issues specific to microbial contaminants in water and water-related media.

The purpose of the risk characterization is to summarize the event of interest according to its nature, severity, and consequences. The risk characterization also frames the risk assessment results within the context of the problem formulation elements, specifically the nature of the concern, the purpose and objectives, the history and context within the agency, and the questions the risk assessment was designed to answer. For example, the risk management options defined during problem formulation can be used to develop risk estimates with and without proposed control measures. A discussion of the most sensitive variables (sensitivity analysis), or the variables with the largest contribution to the overall uncertainty in the risk estimate, may provide risk managers with insights that can be used for future resource allocation for developing risk mitigation strategies. As new data become available or as risk managers ask new questions, the problem formulation and risk assessment can be revisited and revised as needed and appropriate. Discussions of variability, uncertainty, and identified gaps in the knowledge base should be reiterated from the discussions presented in the problem formulation.

Information from the exposure and health effects components of the risk assessment should be integrated to arrive at conclusions for the microbial risk assessment. The key issues that affect the results should be summarized and put into context. For example, the risk characterization includes a discussion and quantifications (to the extent possible) of (1) the uncertainties associated with the analysis and key components; (2) the variability associated with key inputs to the model(s); (3) the confidence in the resulting risk estimates through a weight-of-evidence discussion; (4) the limitations of the analysis; and (5) the plausibility of the results. A candid and open discussion of the uncertainty in the overall assessment, in each of its components, and related estimates of risk is critical to a full characterization of risk. Uncertainty and sensitivity analyses are often conducted to develop the information needed for this purpose (see Section 5.3). As the assumptions, approaches, and conclusions of the risk assessment are presented, the strengths and limitations should also be discussed.

The assessment of data quality should be part of a risk characterization. Whenever possible, the data that are used should be both relevant and of high quality; however, it should be understood that the quality of available information will vary substantially. A candid discussion of the quality of the data employed should be provided, including how the data quality pertains to variability and uncertainty. Sufficient detail should be provided so that the assessment can be duplicated by others. A discussion (at least in a qualitative manner) of how a specific risk compares with similar risks and discussion of the plausibility of the risk scenarios ("ground truthing") is valuable for TCCR. This may be accomplished by comparisons with other pollutants or situations on which the Agency has already decided to act or for other relevant situations. The discussion should highlight the limitations of such comparisons as well as the relevance of the comparisons. Refer to Table 1 for a concise summary of the elements described above.

Risk characterization is essential for managers who are evaluating public-health concerns, considering regulatory and technologic decisions, and setting priorities for research and funding (NRC, 2009). For example, the NRC (2009) identified the following questions as being central for risk characterization: (1) What is the nature and magnitude of risk associated with existing conditions? (2) What risk decreases (benefits) are associated with each of the options? (3) Are any risks increased? and (4) What are the significant uncertainties?

5.1.1. Historical Context

The first significant reference to risk characterization is found in the 1983 NRC publication titled *Risk Assessment in the Federal Government: Managing the Process* (commonly referred to as the "Red Book"). In that seminal work, the NRC defined risk characterization as

> …the process of estimating the incidence of a health effect under the various conditions of human exposure described in exposure assessment. It is performed by combining the exposure and dose-response assessments. The summary effects of the uncertainties in the preceding steps are described in this step.

Since its publication, the concept of risk characterization evolved within EPA and also more broadly within the U.S. Federal government. Concerns over adequately characterizing risk to maintain the public's perception of and confidence in EPA's risk assessments resulted in a 1992 Agency-wide policy for risk characterization, which stated that "...scientific uncertainty is a fact of life (and)...a balanced discussion of reliable conclusions and related uncertainties enhances, rather than detracts, from the overall credibility of each assessment…" (U.S. EPA, 1995a).

In 1997, the Presidential Commission on Risk Assessment and Risk Management noted that "risk characterization is the primary vehicle for communicating health risk assessment findings," but concluded that the difficulty in communicating risk effectively impedes the risk management process (Omenn et al., 1997).

Risk characterization at EPA is considered to be a conscious and deliberate process to bring all important considerations about risk (the likelihood of the risk and also the strengths and limitations of the assessment) and a description of how others have assessed the risk into an integrated picture. Based on the experiences across the Agency between 1995 and 2000, a single Agency-wide document was determined to be needed. The *Risk Characterization Handbook* (U.S. EPA, 2002b) was developed to respond to that need and remains current. However, the *Risk Characterization Handbook* indicates that Agency offices may wish to prepare tailored guidance that meets their individual needs to supplement and remain consistent with the information in the Handbook. This MRA Tools document fills one such need as the field of MRA has evolved rapidly over recent decades.

5.2. Risk Estimation and Risk Description

The risk estimation is the compilation of the types and magnitude of effects anticipated from exposure to the microbe or medium and can be qualitative or quantitative depending on the data and methods used. The risk description involves summarizing the event of interest according to its nature, severity, and consequences. The risk description should also explicitly state whether and how well both the statement of concern, statement of purpose, and objectives that were identified in the problem formulation were addressed. It is this description that is the synthesis of all of the previous components conducted within scope of the assessment.

The results from the exposure assessment (which may have involved a detailed exposure model) can be expressed as the number of organisms to which an individual is exposed in a defined amount

of time and/or for a certain consumption rate, and can include single or repeated exposures. The results from human health effects assessment (which may have involved a dose-response model) can be expressed as the probability of individual infection or illness after a certain number of organisms are ingested. The risk estimation can be expressed as an individual risk estimate (e.g., probability of illness = 0.001) or as a population level risk estimate (100 illnesses per year within a population of 100,000 individuals). As described in further detail in the problem formulation, the risk estimation can also be modeled to consider time-dependent elements such as secondary (person-to-person) transmission, host immunity, and multiple routes of exposure (ILSI, 2000).

5.3. Uncertainty and Sensitivity Analysis

The terms sensitivity, uncertainty, and variability are terms of art in risk assessment, but may also be used in the vernacular by risk managers. Risk assessors should be aware that risk managers might be looking for a more qualitative answer when they ask "how sure are you?" Sensitivity analysis is evaluating which parameter in the risk assessment has the most impact on the results. In other words the results are sensitive to adjustments to the parameter. The parameters that impact the results the most would usually be considered the more important parameters in the risk assessment. For each parameter there is variability, which cannot be reduced by the collection of more data, and uncertainty, which can be reduced by the collection of more data. It is not always possible to separate variability and uncertainty. However, sensitivity analysis can capture both the variability and uncertainty. Variability and uncertainty in the input parameters results in overall uncertainty in the risk assessment output, which is often reflected by the risk estimate being expressed as a range. Risk managers find discussions of uncertainty and sensitivity most useful when they transparently describe the major sources of uncertainty and how significant those sources of uncertainty are with respect to the results of the risk assessment. Risk managers are also interested in which data gaps, if filled, would most improve the risk assessment.

Uncertainty analysis "is the computation of the total uncertainty induced in the output by quantified uncertainties in the inputs and models…" (Morgan and Henrion, 1990). It is a key concern for risk managers because uncertainty analysis provides information about the overall reliability of the risk estimates. Measures of model uncertainty communicate to risk managers the risk assessor's best judgment as to the overall quality of the numerical risk estimates generated by the MRA. Confidence intervals, "credible ranges" developed through Monte-Carlo analyses, Bayesian analyses, and other measures of dispersion in risk estimates, must be presented clearly, and their meaning communicated clearly. Similarly, clear graphical or tabular presentations are very useful. To the extent that intermediate calculations add value and understanding to the results, they can also be included.

Key assumptions related to model selection,[25] input data, and parameters should be provided and discussed, as well as their implications for the model results and uncertainty. In many risk assessments, assumptions and rough estimates for input values and/or uniform and triangular distributions are used to account for uncertainties in input values that cannot be easily quantified. Any conservative assumptions that are built into the model should be explained, and the effect of using less conservative assumptions should be discussed.

[25] One method that has been used to evaluate this source of uncertainty is model averaging (U.S. EPA, 2006a).

Variability and uncertainty can be captured in risk assessments through the use of distributions instead of point estimates. For example, if the input parameters are distributions of values, then the risk assessment software can select a number randomly from each parameter's distribution and calculate risk for that set of numbers, then the random selection of values and calculation of risk is repeated thousands of times. The resulting set of numbers is a distribution itself, and can be considered the risk estimate or output of the risk assessment.

Evaluating the effect of known sources of variability in model outputs can be done through one or more forms of sensitivity analysis. Sensitivity analysis "is the computation of the effect of changes in input values or assumptions (including boundaries and model functional form) on the outputs" (Morgan and Henrion, 1990). These analyses provide an opportunity for assessing the form of the models comprising the QMRA and can be used to identify the parameters to which the risk estimates are most sensitive. Knowledge of the parameters driving the risk estimates provides opportunities for effective risk management. Sensitivity analyses techniques range from simply conducting a small number of additional model runs with different parameter values to performing a fully probabilistic evaluation of the effects of variations in parameter values on model outputs (e.g., using a Monte Carlo approach). The specific approach taken will depend on the nature of the data and models supporting a given assessment.

While sensitivity analyses are useful for evaluating the effects of the variability in single parameters on risk estimates, when multiple parameter values vary, the results of sensitivity analyses must be interpreted cautiously (U.S. EPA, 1997a). If the variations in parameter values are independent of one another, it is easy to overestimate the effect of varying more than one value, because using upper or lower percentile values for more than one variable can yield point estimates of risk that are overly conservative or insufficiently protective. If the variability in risk parameters is correlated, the effect of their variations may not be easy to estimate using sensitivity analysis. In such cases, a more detailed and comprehensive analysis may be required, usually employing probabilistic approaches such as Monte Carlo or related simulation techniques. Where the variability in model parameters can be partitioned into components mainly reflecting variability and uncertainty, "two-dimensional" Monte Carlo analysis can be employed to estimate the relative importance of these two components. (Refer to U.S. EPA, 2006 for an excellent example of a two-dimensional Monte Carlo analysis.) Monte Carlo analysis and the usual "diagnostics" that it generates can also be used both to estimate the overall precision in model outputs and to identify those input parameters that contribute the most to the overall variability in the risk estimates (U.S. EPA, 1997b; FAO/WHO, 2003; Frey et al., 2004). The EPA *Exposures Factors Handbook* (U.S. EPA, 1997a, 2011) provides several approaches to quantitative uncertainty and sensitivity analysis (see Table 11).

Table 11. Approaches to Sensitivity and Uncertainty Analysis Recommended in EPA's Exposure Factors Handbook (Source: U.S. EPA, 1997a)

Approach	Description	Example
Sensitivity analysis	Changing one input variable at a time while leaving others constant to examine affect on output	Fix each input at lower (then upper) bound while holding others at nominal values (e.g., medians)
Analytical uncertainty propagation	Examining how uncertainty in individual parameters affects the overall uncertainty of the exposure assessment	Analytically or numerically obtain a partial derivative of the exposure equation with respect to each input parameter
Probabilistic uncertainty analysis	Varying each of the input variables over various values of their respective probability distributions	Assign probability density function to each parameter; randomly sample values from each distribution and insert into the exposure equation (Monte Carlo simulation)
Classical statistical methods	Estimating the population exposure distribution directly, based on measured values from a representative sample	Compute confidence interval estimates for various percentiles of the exposure distribution

In addition, Morgan and Henrion (1990) discuss in detail four techniques for sensitivity and uncertainty analysis:

- **deterministic**: one-at-a-time analysis of each factor holding all others constant at nominal values;
- **deterministic joint analysis**: changing the value of more than one factor at a time;
- **parametric analysis**: moving one or a few inputs across reasonably selected ranges such as from low to high values in order to examine the shape of the response; and
- **probabilistic analysis**: using correlation, rank correlation, regression, or other means to examine how much of the uncertainty in conclusions is attributable to which inputs.

EPA's *Guiding Principles for Monte Carlo Analysis* (U.S. EPA, 1997b) provides guidance on selecting and developing the conceptual and mathematical models, selecting and evaluating input data and distributions, evaluating variability and uncertainty, and presenting the results of Monte Carlo analysis. In addition to a policy statement for the use of probabilistic analysis in risk assessment at EPA, eight "conditions for acceptance," which are also reflected throughout this MRA document, are outlined and reproduced below:

1. The purpose and scope of the assessment should be clearly articulated in a problem formulation section that includes a full description of any highly exposed or highly susceptible subpopulations evaluated (e.g., children, the elderly). The questions the assessment attempts to answer are to be discussed and the assessment endpoints are to be well defined.
2. The methods used for the analysis (including all models used, all data upon which the assessment is based, and all assumptions that have a significant impact upon the results) are to be documented and easily located in the report. This documentation is to include a discussion of the degree to which data used are representative of the

population under study. Also, this documentation is to include the names of models and software used to generate the analysis. Sufficient information is to be provided to allow the result of the analysis to be independently reproduced.

3. The results of the sensitivity analysis are to be presented and discussed in the report. Probabilistic techniques should be applied to the pathways and factors of importance to the assessment, as determined by sensitivity analyses or other basic requirements of the assessment.
4. The presence or absence of moderate to strong correlations or dependencies between the input variables is to be discussed and accounted for in the analysis, along with the effects these have on the output distribution.
5. Information for each input and output distribution is to be provided in the report. This includes tabular and/or graphical representations of the distributions (e.g., probability density function and cumulative distribution function plots) that indicate the location of any point estimate of interest (e.g., mean, median, 95th percentile). The selection of distributions is to be explained and justified. For both the input and output distributions, variability and uncertainty are to be differentiated where possible.
6. The numerical stability of the central tendency and the higher end (i.e., tail) of the output distributions are to be presented and discussed.
7. Calculations of exposures and risks using deterministic (e.g., point estimate) methods are to be reported if possible and/or appropriate. Providing these values will allow comparisons between the probabilistic analysis and past or screening level risk assessments. Further, deterministic estimates may be used to answer scenario specific questions and to facilitate risk communication. When comparisons are made, it is important to explain similarities and differences in the underlying data, assumptions, and models.
8. Since fixed exposure assumptions (e.g., exposure duration, body weight) are sometimes embedded in the toxicity metrics (e.g., Reference Doses, Reference Concentrations, unit cancer risk factors), the exposure estimates from the probabilistic output distribution are to be aligned with the toxicity metric.[26]

The USDA (Frey et al., 2004) also identified several sensitivity analytical techniques useful for MRA (Table 12). Although the USDA study focused on assessing microbial risks associated with food processing, the general approaches summarized in Table 12 are also applicable to MRAs for other media, including water and water-related media. The methods range from simple and intuitive (varying input values across their observed ranges, scatter plots) to more complex statistical procedures (e.g., classification and regression tree [CART]). For any given risk assessment, it is likely that more than one of these methods will be useful for sensitivity analysis.

[26] Note that acceptance of condition 8 is mainly relevant for chemical risk assessments and might not be relevant for MRA because defaults that apply to all MRAs have not been developed.

Table 12. Sensitivity Analysis Methods and Techniques (Adapted from Frey and Patil, 2002; Frey et al., 2004)

Sensitivity Analysis Type	General Description	Techniques	Description
Mathematical	Quantification of the variation in model output with the range of variation of an input. Typically involves systematic variation of input parameters, evaluation of model, and assessment of the influence of the input parameters on the model output.	Nominal Range Sensitivity Analysis (NRSA)	Variation of individual inputs over their range while holding all other inputs at their nominal values. Sensitivity is assessed via comparison of model outputs corresponding to the range of values. When model output is probability, the difference in log odds ratio (ΔLOR) method may be preferred.
		Differential sensitivity analysis (DSA)	Variation of individual input in small range near central tendency values. Sensitivity is assessed based on variation in model output in the range around the central tendency.
		Automatic differentiation (AD)	This method is similar to DSA, except sensitivity is assessed based on numerical partial derivatives for the variation in model output with changes in input parameters.
		Difference in log odds ratio (ΔLOR)	Similar to NRSA, except sensitivity is assessed via the ΔLOR, where $\Delta\text{LOR} = \ln\left(\frac{p(\text{event} \mid \text{with changes in input})}{p(\text{Not event} \mid \text{with changes in input})}\right) - \ln\left(\frac{p(\text{event} \mid \text{w/out changes in input})}{p(\text{not event} \mid \text{w/out changs in input})}\right)$
		Worst-case determination	Similar to the ΔLOR approach, quantifies sensitivity to a factor via a factor sensitivity ratio given by $FS_k = \log\left(\frac{N_k(\text{extreme})}{N_k(\text{average})}\right)$ where *k* refers to the factor, *N* is the output (e.g., dose in the study conducted by Petterson et al., (2006)) and extreme and average refer to worst-case and baseline values.
		Break-point analysis	Search for values of inputs at which decision-makers would be indifferent between two or more risk management options.
Statistical	Inputs to models are assigned probability distributions and sensitivity is assessed via the effect of variance of the inputs on model output. Inputs may be varied using Monte Carlo simulation, Latin hypercube sampling, or other methods.	Regression techniques (sample regression or rank regression)	Linear models (either based on known relationships or analysis of scatter plots, etc.) are developed for the dependence model output on input variables. Regression is performed on a sample of data generated from the model (e.g., by Latin hypercube sampling, as demonstrated by de Vos et al. (2006). Sensitivity to input variables may be assessed via comparison of standard errors of regression coefficients or via application of stepwise regression techniques.
		Analysis of variance (ANOVA)	ANOVA is used to determine whether there is a statistical relationship between input variables and model output; in contrast to regression techniques, no functional form for the relationship is assumed and data may be qualitative or quantitative.

Table 12. Sensitivity Analysis Methods and Techniques (Adapted from Frey and Patil, 2002; Frey et al., 2004) (continued)

Sensitivity Analysis Type	General Description	Techniques	Description
		Sample (Pearson) correlation or rank (Spearman) correlation	Sample correlation measures the strength of linear association between input variables and model outputs. Rank correlation is a measure of the strength of the monotonic relationship between two random variables.
		Classification and regression tree (CART)	Nonparametric technique that can select from among a large number of variables those and their interactions that are most important in determining the whether an outcome variable reaches a criterion value (Soller and Eisenberg, 2008). Output variables are divided into classes (e.g., above and below a criterion) and a tree of events leading to the output variable is developed and analyzed.
Graphical	Techniques for visualizing the change in model outputs with changes in model parameters.	Scatter plots	Plots providing information on the relationship between input variables and model outputs are constructed.
		Conditional sensitivity analysis (CSA)	Evaluating (usually graphically) the effect of changes in a subset of model inputs while other inputs are held at fixed values.

Additional EPA references relevant to uncertainty analysis include *Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments* (U.S. EPA, 1999c); *Guidelines for Preparing Economic Analyses* (U.S. EPA, 2000e); *Using Probabilistic Methods to Enhance the Role of Risk Analysis in Decision-making with Case Study Examples* (U.S. EPA, 2009b); and the interagency *Microbial Risk Assessment Guideline* (U.S. EPA/USDA, 2012) (Text Box 10).

Text Box 10. Two-Dimensional Probabilistic Risk Analysis of Cryptosporidium in Public Water Supplies, with Bayesian Approaches to Uncertainty Analysis (Source: from case study #8 in U.S. EPA, 2009b)

Probabilistic assessment of the occurrence and health effects associated with *Cryptosporidium* bacteria in public drinking water supplies was used to support the economic analysis of the final LT2 Enhanced Surface Water Treatment Rule. U.S. EPA's Office of Ground Water and Drinking Water conducted this analysis and established a baseline disease burden attributable to *Cryptosporidium* in Public Water supplies that use surface water sources. Next, it models the source water monitoring and finished water improvements that will be realized as a result of the Rule. Post-LT2 Rule risk is estimated and the Rule's health benefit is the result of subtracting this from the baseline disease burden.

Probabilistic Risk Analysis. Probabilistic assessment was used for this analysis as a means of addressing the variability in the occurrence of *Cryptosporidium* in raw water supplies, the variability in the treatment efficiency, as well as the uncertainty in these inputs and in the dose-response relationship for *Cryptosporidium* infection. This case study provides an example of a probabilistic risk assessment that evaluates both variability and uncertainty at the same time and is referred to as a two-dimensional probabilistic risk assessment. The analysis also included probabilistic treatments of uncertain dose-response and occurrence parameters. Markov Chain Monte Carlo samples of parameter sets filled this function. This Bayesian approach (treating the unknown parameters as random variables) differs from classical treatments, which would regard the parameters as unknown, but fixed. The risk assessment used existing datasets (e.g., occurrence of *Cryptosporidium* and treatment efficacy) to inform the variability of these inputs. Uncertainty distributions were characterized based on professional judgment or by analyzing data using Bayesian statistical techniques.

Text Box 10. Two-Dimensional Probabilistic Risk Analysis of Cryptosporidium in Public Water Supplies, with Bayesian Approaches to Uncertainty Analysis (Source: from case study #8 in U.S. EPA, 2009b)

(continued)
(Results of Analysis. The risk assessment identified the *Cryptosporidium* dose-response relationship as the most critical model parameters in the assessment, followed by the occurrence of the pathogen and treatment efficiency. By simulating implementation of the Rule using imprecise, biased measurement methods, the assessment provided estimates of the number of public water supply systems that would require corrective action and the nature of the actions likely to be implemented. This information afforded a realistic measure of the benefits (in reduced disease burden) expected with the LT2 Rule. In response to U.S. EPA's Science Advisory Board comments, additional *Cryptosporidium* dose-response models were added to more fully reflect uncertainty in this element of the assessment.

5.4. Representative Examples of MRAs

There are numerous examples in the literature of well-performed MRAs that have been conducted for a wide range of purposes. For example, MRAs have been conducted by U.S. governmental agencies, various governmental and non-governmental agencies outside of the United States, including the WHO, as well as by researchers investigating local, regional, and national scale issues. It is not feasible to describe all of these MRAs, however, several examples are highlighted below (along with citations) for interested readers.

- EPA used MRA during the development of the Interim Enhanced Surface Water Treatment Rule and the Long Term 2 Enhanced Surface Water Treatment Rule (U.S. EPA, 2002e, 2006a). In both cases, exposure to *Cryptosporidium* spp. through drinking water was the focus of the assessment.
- EPA used MRA to quantify the benefits of the Ground Water Rule (U.S. EPA, 2006b). In this case, the MRA was conducted on viral agents. A static risk model was used to quantify the benefits of the rule and a dynamic model was used to evaluate the potential implications of person-to-person transmission.
- FDA conducted a quantitative assessment of relative risk to public health from foodborne *Listeria monocytogenes* among selected categories of ready-to-eat foods (FDA/USDA, 2003).
- FDA conducted a quantitative risk assessment on the public health effects of pathogenic *Vibrio parahaemolyticus* in raw oysters (FDA, 2005).
- USDA conducted a quantitative risk assessment of *Clostridium perfringens* in ready-to-eat and partially cooked meat and poultry products. The purpose of the risk assessment was to (1) evaluate the public health effect of changing the allowed maximal growth of *C. perfringens* during manufacturing stabilization (cooling after the cooking step) of ready-to-eat and partially cooked meat and poultry products; and (2) examine whether steps taken to limit the growth of *C. perfringens* occurring in ready-to-eat and partially cooked foods would be adequate to protect against growth of *C. botulinum* (Golden et al., 2009).
- USDA conducted a comprehensive risk assessment of *Salmonella enterica* serotype *Enteritidis* (*Salmonella Enteritidis*) in December 1996 in response to an increasing number of human illnesses associated with the consumption of shell eggs. The objectives of this risk assessment were to establish the unmitigated risk of foodborne illness from *Salmonella*

Enteritidis, identify and evaluate potential risk reduction strategies, identify data needs, and prioritize future data collection efforts (USDA, 1998).

- MICRORISK is a collaborative research project with an objective to develop and evaluate a harmonized framework for quantitative assessment of the microbiological safety of drinking water in European Union Member States. MRA plays a central role in this research effort. Numerous scientific publications have resulted from this effort (http://www.microrisk.com/publish/cat_index_25.shtml)
- Several MRAs have been conducted by researchers to inform management options for Australian waters. For example, QMRA was used to estimate the reduction of risk encountered by coastal bathers from the commissioning of deepwater ocean outfalls (Ashbolt et al., 1997), and to identify conditions in which an urban freshwater recreational lake should be closed and re-opened based on rainfall and lake level measurements as surrogates for likely sewage impact and lake clearance of fecal contamination (Roser et al., 2006).
- Similar to one of the Australian examples above, MRA was used by the Ministry for the Environment in New Zealand to form the basis of their revised recreational water criteria (NZ MFE, 2003). Researchers have also used MRA to evaluate the potential public health benefits associated with alternative water and wastewater treatment processes (Weir et al. 2011, Soller et al., 2003), to understand the etiologic agents causing illness during recreational water epidemiological results (Soller et al. 2010c), and more comprehensively understand the causes and dynamics of waterborne and foodborne outbreaks that have occurred in the United States (Eisenberg et al., 1998; Seto et al., 2007). EPA used MRA to understand the potential human health-based implications of various sources of fecal contamination to recreational waters (U.S. EPA, 2010).

In addition to the examples listed above, EPA's National Homeland Security Research Center published a *Compendium of Prior and Current Microbial Risk Assessment Methods for Use as a Basis for the Selection, Development, and Testing of a Preliminary Microbial Risk Assessment Framework*. That literature review includes summaries of 135 studies published between 1994 and 2004— 44 related to exposure assessment (oral, inhalation, and dermal), 31 related to dose-response, and 60 related to risk characterization (U.S. EPA, 2007b).

6. References

This list includes the references cited in this document, including appendices (A-C).

Abad, C., Pintó, R.M., Villena, C., Gajardo, R., and Bosch, A. 1997. Astrovirus survival in drinking water. Applied and Environmental Microbiology 63(8):3119-3122.

Abad, C., Villena, C., Guix, S., Caballero, S., Pintó, R.M., and Bosch, A. 2001. The potential role of fomites in the vehicular transmission of human astroviruses. Applied and Environmental Microbiology 67:3904-3907.

Abbaszadegan, M., Stewart, P., and LeChevallier, M.A. 1999. Strategy for detection of viruses in groundwater by PCR. Applied and Environmental Microbiology 65:444-449.

Abbaszadegan, M. 2006. Rotaviruses. Pp. 295-298 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Aitken, C., and Jeffries, D.J. 2001. Nosocomial spread of viral disease. Clinical Microbiology Reviews 14(3):528-546.

Akin, E.W. 1981. Paper presented at the US EPA symposium on microbial health considerations of soil disposal of domestic wastewaters.

Al-Saleem, T., and Al-Mondhiry, H. 2005. Immunoproliferative small intestinal disease (IPSID): a model for mature B-cell neoplasms. Blood 105(6):2274-2280.

Allen, L.J.S., and Allen, E.A. 2003. A comparison of three different stochastic population models with regard to persistence time. Theoretical Population Biology 64:439-449.

Anderson, R.M. and May, R. 1991. Infectious diseases of humans: Dynamics and Control. New York: Oxford University Press.

Anderson, B.S., Sims, J.K., Liang, A.P., and Minette, H.P. 1988. Outbreak of eye and respiratory irritation in Lahaina, Maui, possibly associated with *Microcoleus lyngbyaceus*. Journal of Environmental Health 50(4):205-209.

Ang, C.W., Noordzij, P.G., de Klerk, M.A., Endtz, H.P., van Doorn, P.A., and Laman, J.D. 2002. Ganglioside mimicry of *Campylobacter jejuni* lipopolysaccharides determines antiganglioside specificity in rabbits. Infection and Immunity 70(9):5081-5085.

Anonymous. 1998. Case Records of the Massachusetts General Hospital. Case 19-1998. New England Journal of Medicine. 338:1830-1836.

Ansari, S.A., Springhorpe, V.S., and Sattar, S.A. 1991. Survival and vehicular spread of human rotaviruses: possible relation to seasonality of outbreaks. Reviews of Infectious Diseases 13:448-461.

APHA (American Public Health Association). 2004. Control of Communicable Diseases Manual. 18th Edition. Washington, DC: APHA.

Arguin, P.M., Kozarsky, P.E., and Reed, C. 2008. Hepatitis, Viral, Type E. Chapter 4 in Prevention of Specific Infectious Diseases, CDC Health Information for International Travel 2008 (Yellow Book). U.S. Centers for Disease Control and Prevention.

Armstrong,T.W., and Haas, C.N. 2008. Legionnaires' disease: evaluation of a quantitative microbial risk assessment model. Journal of Water and Health 6(2):149-166.

Artigas, J., Grosse, G., Niedobitek, F., Kassner, M., Risch, W., and Heise, W. 1994. Severe toxoplasmic ventriculomeningoencephalomyelitis in two AIDS patients following treatment of cerebral toxoplasmic granuloma. Clinical Neuropathology 13(3):120-126.

Asano, T., Leong, L.Y.C., Rigby, M.G., and Sakaji, R.H. 1992. Evaluation of the California wastewater reclamation criteria using enteric virus monitoring data. Water Science and Technology 26(7-8):1513-1524.

Ashbolt, N.J., Reidy, C., and Haas, C.N. 1997. Microbial health risk at Sydney's coastal bathing beaches. In: Proceeding of 17th Australian Water and Wastewater Association meeting. AWWA. Melbourne, pp. 104-111.

ASM (American Society for Microbiology). 2011. Manual of Clinical Microbiology, 10th Edition. Editor in Chief: James Versalovic. http://mcm10.asmpress.org/

AWWA (American Water Works Association). 2006. Waterborne Pathogens, AWWA Manual M48, Second Edition. American Water Works Association: Denver, CO.

Awwad, S.T., Petroll, W.M., McCulley, J.P., and Cavanagh, H.D. 2007. Updates in *Acanthamoeba* keratitis. Eye Contact Lens 33(1):1-8.

Azevedo, N.F., Pinto, A.R., Reis, N.M., Vieira, M.J., and Keevil, C.W. 2006. Shear stress, temperature, and inoculation concentration influence the adhesion of water-stressed *Helicobacter pylori* to stainless steel 304 and polypropylene. Applied and Environmental Microbiology 72(4):2936-2941.

Bailey, N.T.J. 1964. The Elements of Stochastic Processes with Application to the Natural Sciences. New York: John Wiley and Sons.

Balbus, J., Parkin, R., and Embrey, M. 2000. Susceptibility in microbial risk assessment: definitions and research needs. Environmental Health Perspectives 108:901-905.

Ballweber, L.R., Xiao, L., Bowman, D.D., Kahn, G., and Cama, V.A. 2010. Giardiasis in dogs and cats: update on epidemiology and public health significance. Trends in Parasitology 26:180-189.

Barnett, N.D.P., Kaplan, A.M., Hopkin, R.J., Saubolle, M.A., and Rudinsky, M.F. 1996. Primary amoebic meningoencephalitis with *Naegleria fowleri*: clinical review. Pediatric Neurology 15(3): 230-234.

Barrett-Connor, E., and Connor, J.D. 1970. Extraintestinal manifestations of shigellosis. American Journal of Gastroenterology 53(3):234-235.

Bartram, J., Corrales, L., Davison, A., Deere, D., Drury, D., Gordon, B., Howard, G., Rinehold, A., and Stevens, M. 2009. Water Safety Plan Manual: Step-by-Step Risk Management for Drinking-Water Suppliers. Geneva, Switzerland: WHO.

Bartrand, T.A., Weir, M.H., and Haas, C.N. 2008. Dose-Response Models for Inhalation of Bacillus anthracis Spores: Interspecies Comparisons. Risk Analysis 28(4):1115-1124.

Behan, P.O., and Bakheit, A.M.O. 1991. Clinical spectrum of postviral fatigue syndrome. British Medical Bulletin 47(4):793-808.

Bell, J.F., Owens, C.R., and Larson, C.L. 1955. Virulence of *Bacterium tularense*. I. A study of the virulence of *Bacterium tularense* in mice, guinea pigs, and rabbits. Journal of Infectious Diseases 97(2):162-167.

Bellack, N.R., Koehoorn, M.W., MacNab, Y.C., and Morshed, M.G. 2006. A conceptual model of water's role as a reservoir in Helicobacter pylori transmission: a review of the evidence. Epidemiology and Infection 134(3):439-49.

Benenson A.S. 1995. Giardiasis. Control of Communicable Disease in Man. 16th edition, American Public Health Association, Washington, DC.

Betancourt, W.Q., and Rose, J.B. 2004. Drinking water treatment processes for removal of C*ryptosporidium* and *Giardia*. Veterniary Parasitology. 126:219-234.

Bieber, D., Ramer, S.W., Wu, C.Y., Murray, W.J., Tobe, T., Fernandez, R., and Schoolnik, G.K. 1998., Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. Science 280(5372):2114-2118.

Blaser, M.J., and Kirschner, D. 1999. Dynamics of *Helicobacter pylori* colonization in relation to the host response. Proceedings of the National Academy of Sciences (USA) 96(15):8359-8364.

Blaser, M.J., and Kirschner, D. 2007. The equilibria that allow bacterial persistence in human hosts. Nature 449(7164):843-849.

Boehm, A.B., Grant, S.B., Kim, J.H., Mowbray, S.L., McGee, C.D., Clark, C.D., Foley, D.M., and Wellman, D.E. 2002. Decadal and shorter period variability of surf zone water quality at Huntington Beach, California. Environmental Science & Technology 36(18)3885-3892.

Boehm, A.B. 2007. Enterococci concentrations in diverse coastal environments exhibit extreme variability. Environmental Science & Technology 41:8227-8232.

Boerlijst, M.C., Bonhoeffer, S., and Nowak, M.A. 1996. Viral quasi-species and recombination. Proceedings: Biological Sciences 263(1376):1577-1584.

Bogosian, B.J., and Bourneuf, E.V. 2001. A matter of bacteria life and death. EMBO Reports 2(9):770-774.

Bollaerts, K., Aerts, M., Faes, C., Grijspeerdt, K., Dewulf, J., and Mintiens, K. 2008. Human salmonellosis: estimation of dose-illness from outbreak data. Risk Analysis 28(2):427-440.

Borchardt, M.A., Bertz, P.D., Spencer, S.K., and Battigelli, D.A. 2003. Incidence of enteric viruses in groundwater from household wells in Wisconsin. Applied and Environmental Microbiology 69:1172-1180.

Borchardt, M.A., Haas, N.L., and Hunt, R.J. 2004. Vulnerability of drinking-water wells in La Crosse, Wisconsin, to enteric-virus contamination from surface water contributions. Applied and Environmental Microbiology 70:5937-5946.

Boshuizen, H.C., Neppelenbroek, S.E., van Vliet, H., Schellekens, J.F., den Boer, J.W., Peeters, M.F., and Conyn-van Spaendonck, M.A. 2001. Subclinical *Legionella* infection in workers near the source of a large outbreak of legionnaires disease. Journal of infectious Diseases 184:515-518.

Brookmeyer, R., Johnson, E., and Barry, S. 2005. Modelling the incubation period of anthrax. Statistics in Medicine 24:531-542.

Brown, L.M. 2000. *Helicobacter pylori*: epidemiology and routes of transmission. Epidemiologic Reviews 22(2):283-297.

Brown, M.R., and Barker, J. 1999. Unexplored reservoirs of pathogenic bacteria: protozoa and biofilms. Trends in Microbiology 7(1):46-50.

Brunner, R.L., O'Grady, D.J., Partin, J.C., Partin, J.S., and Schubert, W.K. 1979. Neuropsychologic consequences of Reye syndrome. Journal of Pediatrics 95(5 Pt 1):706-711.

Brynestad, S., Braute, L., Luber, P., and Bartelt, E. 2008. Quantitative microbiological risk assessment of campylobacteriosis cases in the German population due to consumption of chicken prepared in homes. International Journal of Risk Assessment and Management 8(3):194-213.

Buchanan, R.L., Smith, J.L., and Long, W. 2000. Microbial risk assessment: dose-response relations and risk characterization. International Journal of Food Microbiology 58:159-172.

Bunning, V.K., Lindsay, J.A., and Archer, D.L. 1997. Chronic health effects of microbial foodborne disease. World Health Statistics Quarterly 50(1-2):51-56.

Butzler, J.P. 2004. *Campylobacter*, from obscurity to celebrity. Clinical Microbiology and Infection 10(10):868-876.

CAC (Codex Alimentarius Commission). 1999. Principles and Guidelines for the Conduct of Microbiological Risk Assessment. (Step 8 of Codex elaboration) CAC/GL-30(1999).

CAC. 2004. Codex Alimentarius Commission – Procedural Manual Fourteenth Edition, Section III – Working Principles for Risk Analysis for Application in the Framework of the Codex Alimentarius.

CAC. 2007. Principles and Guidelines for the Conduct of Microbiological Risk Management. (Step 8 of Codex elaboration) Alinorm 07/30/13. Report of 38th session of the Codex Committee on Food Hygiene. CAC/GL 63-2007

Cacciò, S. 2005. Molecular epidemiology of human cryptosporidiosis. Parassitologia 47:185-192.

Cali, A. 2006. Microsporidia. Pp. 221-228 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Casemore, D.P., Wright, S.E., and Coop, R.L. 1997. Cryptosporidiosis – human and animal epidemiology. In: Cryptospoidium and Cryptosporidiosis, Fayer R (ed), CRC Press, New York.

Castro-Hermida, J.A., García-Presedo, I., Almeida, A., González-Warleta, M., Correia Da Costa, J.M., and Mezo, M. 2008. Contribution of treated wastewater to the contamination of recreational river areas with *Cryptosporidium* spp. and *Giardia duodenalis*. Water Research 42:3528-3538.

CDC (U.S. Centers for Disease Control and Prevention). 1992. Principles of Epidemiology, Self-Study Course 3030-G, Second Edition.
http://www.uic.edu/sph/prepare/courses/ph490/resources/epiintro.pdf.
http://www.uic.edu/sph/prepare/courses/ph490/resources/epilesson01.pdf.

CDC. 1993. Surveillance for Waterborne Disease Outbreaks - United States, 1991-1992. Morbidity and Mortality Weekly Report, 42: 1-22.

CDC. 1996. Surveillance for Waterborne-Disease Outbreaks - United States, 1993-1994. Morbidity and Mortality Weekly Report, 45: 1-33.

CDC. 1998a. *H. pylori:* Fact Sheet for Health Care Providers.
http://www.cdc.gov/ulcer/files/hpfacts.PDF.

CDC. 1998b. Surveillance for Waterborne-Disease Outbreaks - United States, 1995-1996. Morbidity and Mortality Weekly Report, 47: 1-33.

CDC. 2000. Surveillance for Waterborne Disease Outbreaks - United States, 1997-1998. Morbidity and Mortality Weekly Report, 49: 1-35.

CDC. 2002. Surveillance for Waterborne-Disease Outbreaks - United States, 1999-2000. Morbidity and Mortality Weekly Report, 51: 1-48.

CDC. 2004a. About Cyanobacteria. http://www.cdc.gov/hab/cyanobacteria/pdfs/about.pdf.

CDC. 2004b. Centers for Disease Control and Prevention Surveillance for Waterborne-Disease Outbreaks associated with recreational water – United States, 2001-2002 and, Surveillance for Waterborne-Disease Outbreaks associated with drinking water – United States, 2001-2002. Surveillance Summaries, October 22, 2004. Morbidity and Mortality Weekly Reports 53(SS-8).

CDC. 2006a. Non-Polio Enterovirus Infections Website, Division of Viral Diseases. http://www.cdc.gov/ncidod/dvrd/revb/enterovirus/non-polio_entero.htm.

CDC. 2006b. Surveillance for Waterborne Disease and Outbreaks Associated with Recreational Water - United States, 2003-2004. Morbidity and Mortality Weekly Report, 55: 1-30.

CDC. 2008. Surveillance for waterborne-disease and outbreaks associated with recreational water use and other aquatic facilities – United States, 2005-2006. Morbidity and Mortality Weekly Reports 57:1-72.

CDC. 2009a. Foodborne Pathogenic Microorganisms and Natural Toxins Handbook: *Plesiomonas shigelloides.* http://www.foodsafety.gov/~mow/chap18.html.

CDC. 2009b. *Norovirus:* Technical Fact Sheet. http://www.cdc.gov/ncidod/dvrd/revb/gastro/norovirus-factsheet.htm.

CDC. 2009c. Viral Hepatitis Website, Division of Viral Hepatitis. http://www.cdc.gov/hepatitis/HepatitisA.htm.

CDC. 2010. Non-Polio Enterovirus Infections Website, Division of Viral Diseases. http://www.cdc.gov/ncidod/dvrd/revb/enterovirus/non-polio_entero.htm.

CDC/MMWR. 2006. Morbidity and Mortality Weekly Report *55* (SS12):31-58.

Cervenka, L. 2007. Survival and inactivation of *Arcobacter* spp., a current status and future prospect. Critical Reviews in Microbiology 33(2):101-8.

Chappell, C.L., Okhuysen, P.C., Sterling, C.R., and DuPont, H.L. 1996. *Cryptosporidium parvum:* intensity of infection and oocyst excretion patterns in healthy volunteers. J. Infect. Dis., 173:232-236.

Chappell, C.L., Okhuysen, P.C., Sterling, C.R., Wang, C., Jakubowski, W., and DuPont, H.L. 1999. Infectivity of *Cryptosporidium parvum* in healthy adults with pre-existing anti-*C. parvum* serum immunoglobulin G. American Journal of Tropical Medicine and Hygiene 60(1):157-164.

Chappell, C.L., Okhuysen, P.C., Langer-Curry. R., Widmer, G., Akiyoshi, D.E., Tanriverdi, S., and Tzipori, S. 2006. *Cryptosporidium hominis*: experimental challenge of healthy adults. The American Journal of Tropical Medicine and Hygiene 75(5):851-7.

Chen, K.T., Chen ,C.J., and Chiu, J.P. 2001. A school waterborne outbreak involving both *Shigella sonnei* and *Entamoeba histolytica*. Journal of Environmental Health 64(4):9-13, 26.

Chen, T.L., Chan, C.C., Chen, H.P., Fung, C.P., Lin, C.P., Chan, W.L., and Liu, C.Y. 2003. Clinical characteristics and endoscopic findings associated with *Blastocystis hominis* in healthy adults. American Journal of Tropical Medicine and Hygiene 69:213-216.

Chey, W.D., and Wong, B.C.Y. 2007. American College of Gastroenterology guideline on the management of *Helicobacter pylori* infection. American Journal of Gastroenterology 102:1808-1825.

Codd, G.A. 2000. Cyanobacterial toxins, the perception of water quality, and the prioritisation of eutrophication control. Ecological Engineering 16(1):51-60.

Codd, G.A., Ward, C.J., and Bell, S.G. 1997. Cyanobacterial toxins: occurrence, modes of action, health effects and exposure routes. Archives of Toxicology Supplement 19:399-410.

Coleman, M., and Marks, H. 1998. Topics in dose-response modeling. Journal of Food Protection 61:1550-1559.

Coleman, M., and Marks, H. 2000. Mechanistic modeling of salmonellosis. Quantitative Microbiology 2:227-247.

Coleman, M.E., Marks, H.M., Golden, N.J., and Latimer, H.K. 2004. Discerning strain effects in microbial dose-response data. Journal of Toxicology and Environmental Health-Part A –Current Issues 67:667-685.

Colwell, R.R., Huq, A., Islam, M.S., Aziz, K.M.A., Yunus, M., Khan, N.H., Mahmud, A., Sack, R.B., Nair, G.B., Chakraborty, J., Sack, D.A., and Russek-Cohen, E. 2003. Reduction of cholera in Bangladeshi villages by simple filtration. Proceedings of the National Academy of Sciences (USA) 100(3):1051-1055.

Connor, B.A., Johnson, E.J., and Soave, R. 2001. Reiter syndrome following protracted symptoms of *Cyclospora* infection. Emerging Infectious Diseases *7*(3):453-454.

Couch, R.B., Cate, T.R., Gerone, P.J., Fleet, W.F., Lang, D.J., Griffith, W.R., and Knight, V. 1965. Production of illness with a small-particle aerosol of coxsackie A21. Journal of Clinical Investigation 44:535-542.

Covacci, A., and Rappuoli, R. 1998. *Helicobacter pylori*: molecular evolution of a bacterial quasi-species. Current Opinion in Microbiology 1(1):96-102.

Covert, T.C. 1999. *Salmonella*. In Waterborne Pathogens, AWWA manual M48, 1st Edition. Denver, CO: American Water Works Association.

Covert, T.C., and Meckes, M.C. 2006. *Salmonella*. Pp. 135-140 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Crabtree, K.D., Gerba, C.P., Rose, J.B., and Haas, C.N. 1997. Waterborne adenovirus: a risk assessment. Water Science and Technology 35(11-12):1-6.

Crainiceanu, C.M., Stedinger, J.R., Ruppert, D., and Behr, C.T. 2003. Modeling the United States national distribution of waterborne pathogen concentrations with application to *Cryptosporidium parvum*. Water Resources Research 39(9):SWC 2-1.

Craun, M.F., Craun, G.F., Calderon, R.L., and Beach, M.J. 2006. Waterborne outbreaks reported in the United States. Journal of Water and Health 4(Suppl 2):19-30.

Craun, G.F., Brunkard, J.M., Yoder, J.S., Roberts, V.A., Carpenter, J., Wade, T., Calderon, R.L., Roberts, J.M., Beach, M.J. and S.L. Roy. 2010. Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. Clinical Microbiological Reviews, 23(3):507-528.

Davidson, P.W., Otuama, L.A., Willoughby, R.H., and Swisher, C.N. 1978. Neurological and intellectual sequelae of Reye's syndrome. American Journal of Mental Deficiency 82(6): 535-541.

de Vos, C.J., Saatkamp, H.W., Nielen, M., and Huirne1, R.B.M. 2006. Sensitivity Analysis to Evaluate the Impact of Uncertain Factors in a Scenario Tree Model for Classical Swine Fever Introduction. Risk Analysis 26(5):1311-1322.

Didier, E.S., and Weiss, L.M. 2006. Microsporidiosis: current status. Current Opinion in Infectious Diseases 19(5):485-492.

Didier, E.S., Stovall, M.E., Green, L.C., Brindley, P.J., Sestak, K., and Didier, P.J. 2004. Epidemiology of microsporidiosis: sources and modes of transmission. Veterniary Parasitology 126(1-2):145-166.

Dietz, V., Vugia, D., Nelson, R., Wicklund, J., Nadle, J., McCombs, K.G., and Reddy, S. 2000. Active, multisite, laboratory-based surveillance for *Cryptosporidium parvum*. American Journal of Tropical Medicine and Hygiene 62:368-372.

Dorevitch, S., Panthi, S., Huang, Y., Li, H., Michalek, A.M., Pratap, P., Wroblewski, M., Liu, L., Scheff, P.A., and Li, A. 2011. Water ingestion during water recreation. Water Research 45(5): 2020-2028.

Dubey, J.P. 2004. Toxoplasmosis – a waterborne zoonosis. Veterinary Parasitology 126(1-2):57-72.

Dubey, J.P. 2006. *Toxoplasma gondii*. Pp. 239-242 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Dufour, A., Evans, O., Behymer, T., and Cantu, R. 2006. Water ingestion during swimming activities in a pool: a pilot study. Journal of Water and Health 4:425-430.

DuPont, H., Chappell, C., Sterling, C., Okhuysen, P., Rose, J., and Jakubowski, W. 1995. The infectivity of *Cryptosporidium parvum* in healthy volunteers. New England Journal of Medicine 332:855-859.

Dworkin, M.S., Shoemaker, P.C., Goldoft, M.J., and Kobayashi, J.M. 2001. Reactive arthritis and Reiter's syndrome following an outbreak of gastroenteritis caused by *Salmonella enteritidis*. Clinical Infectious Diseases 33(7):1010-1014.

Effler, P., Ieong, M-C., Kimura, A., Nakata, M., Burr, R., Cremer, E., and Slutsker, L. 2001. Sporadic *Campylobacter jejuni* infections in Hawaii: associations with prior antibiotic use and commercially prepared chicken. Journal of Infectious Diseases 83:1152-1155.

Eisenberg, J.N., Seto, E.Y.W., Olivieri, A.W., and Spear, R.C. 1996. Quantifying water pathogen risk in an epidemiological framework. Risk Analysis 16:549-563.

Eisenberg, J.N., Seto, E.Y.W., Colford, J.M., Olivieri, A.W., and Spear, R.C. 1998. An analysis of the Milwaukee cryptosporidiosis outbreak based on a dynamic model of the infection process. Epidemiology 9:255-263.

Eisenberg, J.N., Brookhart, M.A., Rice, G., Brown, M., and Colford, J.M., Jr. 2002. Disease transmission models for public health decision making: analysis of epidemic and endemic conditions caused by waterborne pathogens. Environmental Health Perspectives 110(8):783-790.

Eisenberg, J.N.S., Soller, J.A., Scott, J., Eisenberg, D.M., and Colford, J.M. 2004. A dynamic model to assess microbial health risks associated with beneficial uses of biosolids. Risk Analysis 24:221-236.

Eisenberg, J.N.S., Lei, X., Hubbard, A.H., Brookhart, M.A., and Colford, J.M. 2005. The role of disease transmission and conferred immunity in outbreaks: analysis of the 1993 *Cryptosporidium* outbreak in Milwaukee, Wisconsin. American Journal of Epidemiology 16(1):62-72.

Eisenberg, J.N., Moore, K., Soller, J.A., Eisenberg, D., and Colford, J.M., Jr. 2008. Microbial risk assessment framework for exposure to amended sludge projects. Environmental Health Perspectives 116:727-733.

Ekwall, E., Ljungh, A., and Selander, B. 1984. Asymptomatic urinary tract infection caused by *Shigella sonnei*. Scandinavian Journal of Infectious Diseases 16(1):121-122.

Embrey, M.A. 2002. Cyanobacteria in Drinking Water. In Handbook of CCL Microbes in Drinking Water (pp. 203-227). Denver, CO: American Water Works Association.

Englehardt, J.D. 2004. Predictive Bayesian dose-response assessment for appraising absolute health risk from available information. Human and Ecological Risk Assessment 10(1):69-78.

Englehardt, J.D., and Swartout, J. 2004. Predictive population dose-response assessment for *Cryptosporidium parvum*: infection endpoint. Journal of Toxicology and Environmental Health-Part A-Current Issues 67(8-10):651-666.

Englehardt, J.D, and Swartout, J. 2006. Predictive Bayesian microbial dose-response assessment based on suggested self-organization in primary illness response: *Cryptosporidium parvum*. Risk Analysis 26(2):543-554.

Englehardt, J.D., and Swartout, J. 2008. Development and Evaluation of Novel Dose-Response Models for Use in Microbial Risk Assessment, Technical Report. EPA/600/R-08/033. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency.

Enriquez, C., and Thurston-Enriquez, J. 2006. Adenovirus. Pp. 253-248 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

EO (Executive Order). 1994. Executive Order 12898—Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations. http://www.epa.gov/fedrgstr/eo/eo12898.htm.

EO. 1997. Executive Order 13045—Protection of Children from Environmental Health Risks and Safety Risks. Federal Register 62(78):19883-19888.

Ernst, P.B., and Gold, B.D. 2000. The disease spectrum of *Helicobacter pylori*: the immunopathogenesis of gastroduodenal ulcer and gastric cancer. Annual Review of Microbiology 54:615-640.

Fan, P.C., Chen, Y.C., Tian, Y.C., Chang, C.H., Fang, J.T., and Yang, C.W. 2009. Acute renal failure associated with acute non-fulminant Hepatitis A: a case report and review of literature. Renal Failure 31(8):756-764.

Fankhauser, R.L., Noel, J.S., Monroe, S.S., Ando, T., and Glass, R.I. 1998. Molecular epidemiology of "Norwalk-like viruses" in outbreaks of gastroenteritis in the United States. Journal of Infectious Diseases 178:1571-1578.

FAO/WHO (Food and Agriculture Organization of the United Nations/World Health Organization). 2003. Microbiological Risk Assessment Series, No. 3: Hazard Characterization for Pathogens in Food and Water, Guidelines. http://whqlibdoc.who.int/publications/2003/9241562374.pdf.

FAO/WHO. 2009. Risk Characterization of Microbiological Hazards in Food: Guidelines, MRA Series 17, http://www.who.int/foodsafety/publications/micro/MRA17.pdf

Farber, J.M., Ross, W.H., Harwig, J. 1996. Health risk assessment of *Listeria monocytogenes* in Canada. International Journal of Food Microbiology 30(1-2):145-156.

Faruque, S.M., Biswas, K., Udden, S.M.N., Ahmad, Q.S., Sack, D.A., Nair, G.B., and Mekalanos, J.J. 2006. Transmissibility of cholera: in vivo-formed biofilms and their relationship to infectivity and persistence in the environment. Proceedings of the National Academy of Sciences (USA) 103(16):6350-6355.

Fayer, R. 2004. *Cryptosporidium*: a waterborne zoonotic parasite. Veterinary Parasitology. 126:37-56.

Fayer, R., Speer, C.A., and Dubey, J.P. 1997. The General Biology of *Cryptosporidium*. In *Cryptosporidium* and Cryptosporidiosis, ed. R. Fayer. New York: CRC Press.

Fazil, A., Paoli, G., Lammerding, A.M., Davidson, V., Hrudey, S., Isaac-Renton, J., and Griffiths, M. 2005. Microbial Risk Assessment as a Foundation for Informed Decision-Making: A Needs, Gaps and Opportunities Assessment (NGOA) for Microbial Risk Assessment in Food and Water. Public Health Agency of Canada. http://www.uoguelph.ca/crifs/NGOA/Finalupdates/NGOAfinalreport.pdf.

FDA (U.S. Food and Drug Administration). 2002. Initiation and Conduct of all "Major" Risk Assessments Within a Risk Analysis Framework. A Report by the Center for Food Safety and Applied Nutrition Risk Analysis Working Group. http://www.cfsan.fda.gov/~dms/rafw-toc.html.

FDA/USDA. 2003. Quantitative Assessment of Relative Risk to Public Health from Foodborne Listeria monocytogenes Among Selected Categories of Ready-to-Eat Foods. Rockville, MD. http://www.foodsafety.gov/~dms/lmr2-toc.html.

FDA. 2005. Quantitative Risk Assessment on the Public Health Impact of Vibrio parahaemolyticus in Raw Oysters. Center for Food Safety and Applied Nutrition. http://www.fda.gov/Food/ScienceResearch/ResearchAreas/RiskAssessmentSafetyAssessment/ucm050421.htm.

FDA. 2006. The Bad Bug Book: Foodborne Pathogenic Organisms and Natural Toxins Handbook. Washington, DC: FDA. http://www.fda.gov/food/foodsafety/foodborneillness/foodborneillnessfoodbornepathogensnaturaltoxins/badbugbook/default.htm.

FDA. 2009. Bad Bug Book: Foodborne Pathogenic Microorganisms and Natural Toxins Handbook *Campylobacter jejuni.* http://www.fda.gov/Food/FoodSafety/FoodborneIllness/FoodborneIllnessFoodbornePathogensNaturalToxins/BadBugBook/ucm070024.htm.

Feazel, L.M., Baumgartnera, L.K., Petersona, K.L., Franka, D.N., Harris, J.K., and Pace, N.R. 2009. Opportunistic pathogens enriched in showerhead biofilms. Proceedings of the National Academy of Sciences (USA) 106(38):16393-16399.

Ferrera, P.C., Jeanjaquet, M.S., and Mayer, D.M. 1996. Shigella-induced encephalopathy in an adult. The American Journal of Emergency Medicine 14(2):173-5.

Field, S.K., Fisher, D., and Cowie, R.L. 2004. *Mycobacterium avium* complex pulmonary disease in patients without HIV infection. Chest 126(2):566-581.

Fields, B.S., Benson, R.F., and Besser, R.E. 2002. *Legionella* and Legionnaires' disease: 25 years of investigation. Clinical Microbiology Reviews 15(3):506-526.

Fine, K.D., and Stone, M.J. 1999. Alpha-heavy chain disease, Mediterranean lymphoma, and immunoproliferative small intestinal disease – a review of clinicopathological features, pathogenesis, and differential diagnosis. American Journal of Gastroenterology 94(5): 1139-1152.

Fink, J.N., Ortega, H.G., Reynolds, H.Y., Cormier, Y.F., Fan, L.L., Franks, T.J., Kreiss, K., Kunkel, S., Lynch, D., Quirce, S., Rose, C., Schleimer, R.P., Schuyler, M.R., Selman, M., Trout, D., Yoshizawa, Y. 2005. Needs and opportunities for research in hypersensitivity pneumonitis. American Journal of Respiratory and Critical Care Medicine 171(7):792-798.

Fohlman, J., and Friman, G. 1993. Is juvenile diabetes a viral disease? Annals of Medicine 25(6):569-574.

Fraser, D.W., Tsai, T.R., Orenstein, W., Parkin, W.E., Beecham, H.J., Sharrar, R.G., Harris, J., Mallison, G.F., Martin, S.M., McDade, J.E., Shepard, C.C., and Brachman, P.S. 1977. Legionnaires' disease: description of an epidemic of pneumonia. New England Journal of Medicine 297:1189-1197.

Frey, H.C., Mokhtari, H., and Zheng, J. 2004. Recommended Practice Regarding Selection, Application, and Interpretation of Sensitivity Analysis Methods Applied to Food Safety Process Risk Models. Office of Risk Assessment and Cost-Benefit Analysis, U.S. Department of Agriculture. http://www.ce.ncsu.edu/risk/Phase3Final.pdf.

Fricker, C.R. 2006. *Yersinia*. Pp. 157-160 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Frost, F. and Craun, G.F. 1998. The Importance of Acquired Immunity in the Epidemiology of Cryptosporidiosis and Giardiasis. EPA, OECD Workshop Molecular Methods for Safe Drinking Water.

Frost, F.J., Roberts, M., Kunde, T.R., Craun, G., Tollestrup, K., Harter, L., and Muller, T. 2005. How clean must our drinking water be: the importance of protective immunity. Journal of Infectious Diseases 191:809-814.

Gale, P. 2003. Using event trees to quantify pathogen levels on root crops from land application of treated sewage sludge. Journal of Applied Microbiology 94:35-47.

Gale, P. 2005. Land application of treated sewage sludge: quantifying pathogen risks from consumption of crops. Journal of Applied Microbiology 98:380-396.

Garcia, L.S. 2006a. *Blastocystis hominis*. Pp. 189-192 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Garcia, L.S. 2006b. *Isospora belli*. Pp. 217-220 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Garg, A.X., Suri, R.S., Barrowman, N., Rehman, F., Matsell, D., Rosas-Arellano, M.P., Salvadori, M., Haynes, R.B., and Clark W.F. 2003. Long-term renal prognosis of diarrhea-associated hemolytic uremic syndrome – a systematic review, meta-analysis, and meta-regression. Journal of the American Medical Association 290(10):1360-1370.

Gerba, C.P. 2006. Hepatitis E Virus. Pp. 279-280 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Gerba, C., Yates, M., Yates, S., and Hurst, C., eds. 1991. Quantitation of Factors Controlling Viral and Microbial Transport in the Subsurface. In: Modeling the Environmental Fate of Microorganisms, Washington, DC: American Society for Microbiology.

Gerba, C.P.; Rose, J.B., and Haas, C.N. 1996a. Sensitive populations: who is at the greatest risk? International Journal of Food Microbiology 30(1-2):113-123.

Gerba, C.P., Rose, J.B., Haas, C.N., and Crabtree, K.D. 1996b. Waterborne rotavirus: a risk assessment. Water Research 30:2929-2940.

Gerba, C.P., Henze, M., Loosdrecht, C.M., Ekman, G.A., and Brdjanovic, D., eds. 2008. Pathogen Removal. In: Biological Wastewater Treatment, Modeling and Design, London, UK: IWA Publishing.

Gilbert, M., Godschalk, P.C., Karwaski, M.F., Ang, C.W., van Belkum, A., Li, J., Wakarchuk, W.W., and Endtz, H.P. 2004. Evidence for acquisition of the lipopolysaccharide biosynthesis locus in *Campylobacter jejuni* GB11, a strain isolated from a patient with Guillain-Barre syndrome, by horizontal exchange. Infection and Immunity 72(2):1162-1165.

Gilinsky, N.H., Novis, B.H., Wright, J.P., Dent, D.M., King, H., and Marks, I.N.S. 1987. Immunoproliferative small-intestinal disease: clinical features and outcome in 30 cases. Medicine 66(6):438-446.

Gilks, W., Richardson, S., and Spiegelhalter D.J. eds. 1996. Markov Chain Monte Carlo in Practice. London, UK: Chapman and Hall.

Gilks, W.R. and Wild, P. 1992. Adaptive rejection sampling for Gibbs sampling. Applied Statistics 41:337-348.

Girschick, H.J., Guilherme, L., Inman, R.D., Latsch, K., Rihl, M., Sherer, Y., Shoenfeld, Y., Zeidler, H., Arienti, S., and Doria, A. 2008. Bacterial triggers and autoimmune rheumatic diseases. Clinical and Experimental Rheumatology 26(1 Suppl 48):S12-S17.

Glass, R. I., Bresee, J., Jiang, B. M., Gentsch, J., Ando, T., Fankhauser, R., Noel, J., Parashar, U., Rosen, B., and Monroe, S.S. 2001. Gastroenteritis viruses: an overview. *Nov*artis Foundation Symposium 238:5-25.

Glick, T.H., Gregg, M.B., Berman, B., Mallison, G., Rhodes ,W.W., Jr., and Kassanoff, I. 1978. Pontiac fever. An epidemic of unknown etiology in a health department. I. Clinical and epidemiologic aspects. American Journal of Epidemiology 107:149-160.

Gold, M.R., Stevenson, D., and Fryback, D.G. 2002. HALYs and QALYs and DALYs, oh my: similarities and differences in summary measures of population health. Annual Review of Public Health 23:115-134.

Golden, N.J., Crouch, E.A., Latimer, H., Kadry, A., and Kause, J. 2009. Risk assessment for *Clostridium perfringens* in ready-to-eat and partially cooked meat and poultry products. Journal of Food Protection 72(7):1376-1384.

Goren, A., Freier, S., and Passwell, J.H. 1992. Lethal toxic encephalopathy due to childhood shigellosis in a developed country. Pediatrics 89(6):1189-1193.

Greenberg, H., Valdesuso, J., Kapikian, A., Chanock, R., Wyatt, R., Szmuness, W., Larrick, J., Kaplan, J., Gilman, R.H., and Sack, D.A. 1979. Prevalence of antibody to the Norwalk virus in various countries. Infection and Immunity 26:270-273.

Grohmann, G.S., Glass, R.I., Pereira, H.G., Monroe, S.S., Hightower, A.W., Weber, R., and Bryan, R.T. 1993. Enteric viruses and diarrhea in HIV-infected patients. Enteric Opportunistic Infections Working Group. New England Journal of Medicine, 329(1):14-20.

Gronewold, A.D., Borsuk, M.E., Wolpert, R.L., and Reckhow, K.H. 2008. An assessment of fecal indicator bacteria-based water quality standards. Environmental Science & Technology 42(13):4676-4682.

Gronewold, A.D., Qian, S.S., Wolpert, R.L., and Reckhow, K.H. 2009. Calibrating and validating bacterial water quality models: a Bayesian approach. Water Research 43:2688-2698.

Guix, S., Bosch, A., and Pintó, R.M. 2005. Human astrovirus diagnosis and typing: current and future prospects. Letters in Applied Microbiology 41(2):103-105.

Gupta, A., Polyak, C.S., Bishop, R.D., Sobel, J., and Mintz, E.D. 2004. Laboratory-confirmed shigellosis in the United States, 1989-2002: epidemiologic trends and patterns. Clinical Infectious Diseases 38(10):1372-1377.

Gutting, B.W., Channel, S.R., Berger, A.E., Gearhart, J.M., Andrews, G.A., and Sherwood, R.L. 2008. Mathematically modeling inhalation anthrax. Microbe 3(2):78-85.

Gyori, E. 2003. December 2002: 19-year old male with febrile illness after jet ski accident. *Brain Pathology* 13(2):237-239.

Haas, C.N. 1983. Effect of effluent disinfection on risks of viral disease transmission via recreational water exposure. Journal – Water Pollution Control Federation 55:1111-1116.

Haas, C.N., Crockett, C.S., Rose, J.B., Gerba, C.P., and Fazil, A.M. 1996. Assessing the risks posed by oocysts in drinking water. Journal of the American Water Works Association 88(9):131-136.

Haas, C.N., Rose, J., and Gerba, C.P. 1999. Quantitative Microbial Risk Assessment. New York: Wiley.

Halbur, P.G., Kasorndorkbua, C., Gilbert, C., Guenette, D., Potters, M.B., Purcell, R.H., Emerson, S.U., Toth, T.E., and Meng, X.J. 2001. Comparative Pathogenesis of Infection of Pigs with Hepatitis E Viruses Recovered from a Pig and a Human, Journal of Clinical Microbiology 39(3):918-923.

Hall, N.H. 2006. *Legionella*. Pp. 119-124 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Hanak, V., Golbin, J.M., and Ryu, J.H. 2007. Causes and presenting features in 85 consecutive patients with hypersensitivity pneumonitis. Mayo Clinic Proceedings 82(7):812-816.

Hänninen, M.L., Haajanen, H., Pummi, T., Wermundsen, K., Katila, M.L., Sarkkinen, H., Miettinen, I., and Rautelin, H. 2003. Detection and typing of *Campylobacter jejuni* and *Campylobacter coli* and analysis of indicator organisms in three waterborne outbreaks in Finland. Applied and Environmental Microbiology 69(3):1391-1396.

Hastings, W.K. 1970. Monte Carlo sampling methods using Markov Chains and their applications. Biometrika 57:97-109.

Hejkal, T.W.,Keswick, B., LaBelle, R.L., Gerba, C.P., Sanchez, Y.,Dreesman, G., Hafkin, B., and Melnick, J.L. 1982. Viruses in a community water supply associated with an outbreak of gastroenteritis and infectious hepatitis. Journal of the American Water Works Association 74:318-321.

Hellard, M.E., Sinclair, M.I., Hogg, G.G., and Fairley, C.K. 2000. Prevalence of enteric pathogens among community based asymptomic individuals. Journal of Gastroenterology and Hepatology 15:290-293.

Herath, G. 1995. The algal bloom problem in Australian waterways: an economic appraisal. Review of Marketing and Agricultural Economics 63:77-86.

Herwaldt, B.L. 2000. *Cyclospora cayetanensis*: a review, focusing on the outbreaks of cyclosporiasis in the 1990s. Clinical Infectious Disease 31(4):1040-1057.

Hoeger, S. J., Hitzfeld, B.C., and Dietrich, D.R. 2005. Occurrence and elimination of cyanobacterial toxins in drinking water treatment plants. Toxicology and Applied Pharmacology 203(3):231-242.

Holcomb, D.L., Smith, M.A., Ware, G.O., Hung, Y.C., Brackett, R.E., and Doyle, M.P. 1999. Comparison of six dose-response models for use with food-borne pathogens. Risk Analysis 19(6):1091-1100.

Holdenfried, R. and Quan, S.F. 1956. Susceptibility of New Mexico rodents to experimental plague. Public Health Reports 71(10):979-984.

Holme, R. 2003. Drinking water contamination in Walkerton, Ontario: positive resolutions from a tragic event. Water Science and Technology 47(3):1-6.

Hopkins, R.S., Gaspard, G.B., Williams, F.P., Karlin, R.J., Cukor, K.G., and Blacklow, N.R. 1984. A community waterborne gastroenteritis outbreak: evidence for rotavirus as the agent. American Journal of Public Health 74:263-265.

Hopkins, R.S., Shillam, P., Gaspard, B., Eisnach, L. and Karlin, R.J. 1985. Waterborne disease in Colorado: three years' surveillance and 18 outbreaks. American Journal of Public Health 75(3):254-257.

Hunter, P.R. 1992. Cyanobacteria and human health. Journal of Medical Microbiology 36(5): 301-302.

Huq, M.I., and Islam, M.R. 1983. Microbiological and clinical studies in diarrhoea due to *Plesiomonas shigelloides*. Indian Journal of Medical Research 77:793-797.

IARC (International Agency for Research on Cancer). 1994. Shistosomes, Liver Flukes, and *Helicobacter pylori*. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans 61:1-241.

Illingworth, C. D., and Cook, S.D. 1998. *Acanthamoeba* keratitis. Survey of Ophthalmology 42(6):493-508.

ILSI (International Life Sciences Institute). 1996. A conceptual framework for assessment of the risks of human disease following exposure to waterborne pathogens. Risk Analysis 16:841-848.

ILSI. 2000. Revised Framework for Microbial Risk Assessment. Washington, DC. http://www.ilsi.org/file/mrabook.pdf.

Jameel, S. 1999. Molecular biology and pathogenesis of hepatitis E virus. Expert Reviews in Molecular Medicine (6 December):1-16. http://www-ermm.cbcu.cam.ac.uk/99001271h.htm.

Jiang, S. 2006. Human adenoviruses in water: occurrence and health implications: a critical review. Environmental Science & Technology 40:7132-7140.

Jones, I.G., and Roworth, M. 1996. An outbreak of *Escherichia coli* O157 and campylobacteriosis associated with contamination of a drinking water supply. Public Health 110(5):277-282.

Kaldor, J., and Speed, B.R. 1984. Guillain-Barré syndrome and *Campylobacter jejuni*: a serological study. British Medical Journal 288(6434):1867-1870.

Kappus K.D., Lundgren R.G., Juranek D.D., Roberts J.M., Spencer H.C. 1994. Intestinal parasitism in the United States: update on a continuing problem. American Journal of Tropical Medicine and Hygiene 50:705-713.

Karanis, P., Kourenti, C., and Smith, H. 2007.Waterborne transmission of protozoan parasites: a worldwide review of outbreaks and lessons learnt. Journal of Water Health 5(1):1-38.

Karanja, R.M., Gatei, W., and Wamae, N. 2007. Cyclosporiasis: an emerging public health concern around the world and in Africa. African Health Sciences 7(2):62-67.

Keene, W. 2006 *Entamoeba histolytica*. Pp. 203-207 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Khan, N.A. 2006. *Acanthamoeba*: biology and increasing importance in human health. *FEMS* Microbiology Reviews 30(4):564-595.

Khan, W., Dhar, U., Salam, M., Griffiths, J., Rand, W., and Bennish, M. 1999. Central nervous system manifestations of childhood shigellosis: prevalence, risk factors, and outcome. Pediatrics 103(2):E18.

Kim, K.S., Hufnagel, G., Chapman, N.M., and Tracy, S. 2001. The group B coxsackieviruses and myocarditis. Reviews in Medical Virology 11:355-368.

Kim, S.K., An, J.Y., Park, M.S., and Kim, B.J. 2007. A case report of Reiter's syndrome with progressive myelopathy. Journal of Clinical Neurology 3(4):215-218.

King, A.A., Ionides, E.L., Pascual, M., and Bouma, M.J. 2008. Inapparent infections and cholera dynamics. Nature 454:877-880.

Kistemann, T., Rind, E., Rechenburg, A., Koch, C., Classen, T., Herbst, S., Wienand, I., and Exner, M. 2008. A comparison of efficiencies of microbiological pollution removal in six sewage treatment plants with different treatment systems. International Journal of Hygiene and Environmental Health 211:534-545.

Ko, G., Cromeans, T. L.,and Sobsey, M.D. 2003. Detection of infectious adenovirus in cell culture by mRNA reverse transcription-PCR. Applied and Environmental Microbiology 69(12):7377-7384.

Kodell, R.L., Kang, S-H., and Chen, J.J. 2002. Statistical models of health risk due to microbial contamination of foods. Environmental and Ecological Statistics 9:259-271.

Koopman, J.S., Monto, A.S., and Longini, I.M., Jr. 1989. The Tecumseh Study XVI: Family and community sources of rotavirus infection. American Journal of Epidemiology 130(4)760-768.

Koopman, J.S., Chick, S.E., Simon, C.P., Riolo, C.S., and Jacquez, G. 2002. Stochastic effects on endemic infection levels of disseminating versus local contacts. Mathematical Biosciences 180:49-71.

Kuipers, E.J. 1999. Exploring the link between *Helicobacter pylori* and gastric cancer. Alimentary Pharmacology & Therapeutics 13(Suppl 1):3-11.

Kuipers, E.J., Thijs, J.C., and Festen, H.P. 1995. The prevalence of *Helicobacter pylori* in peptic ulcer disease. Alimentary Pharmacology & Therapeutics 9(Suppl 2):59-69.

Labine, M.A., and Minuk, G.Y. 2009. Cyanobacterial toxins and liver disease. Canadian Journal of Physiology and Pharmacology 87(10):773-788.

Lacasse, Y., Selman, M., Costabel, U., Dalphin, J.C., Ando, M., Morell, F., Erkinjuntti-Pekkanen, R., Muller, N., Colby, T.V., Schuyler, M., and Cormier, Y. 2003. Clinical diagnosis of hypersensitivity pneumonitis. American Journal of Respiratory and Critical Care Medicine 168(8): 952-958.

Lau, H.Y., and Ashbolt, N.J. 2009. The role of biofilms and protozoa in Legionella pathogenesis: implications for drinking water. Journal of Applied Microbiology 107(2):368-378.

LeChevallier, M.W. 2006. *Mycobacterium avium* Complex. Pp. 125-130 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Lecuit, M., Abachin, E., Martin, A., Poyart, C., Pochart, P., Suarez, F., Bengoufa, D., Feuillard, J., Lavergne, A., Gordon, J.I., Berche, P., Guillevin, L., and Lortholary, O. 2004. Immunoproliferative small intestinal disease associated with *Campylobacter jejuni*. New England Journal of Medicine 350(3):239-248.

Lehner, A., Tasara, T., and Stephan, R. 2005. Relevant aspects of *Arcobacter* spp. as potential foodborne pathogen. International Journal of Food Microbiology 102(2):127-135

Levin, B.R., and Antia, R. 2001. Why we don't get sick: the within-host population dynamics of bacterial infections. Science 292(5519):1112-1115.

Lin, H.C., Kao, C.L., Chang, L.Y., Hsieh, Y.C., Shao, P.L., Lee, P.I., Lu, C.Y., Lee, C.Y., and Huang, L.M. 2008. Astrovirus gastroenteritis in children in Taipei. Journal of the Formosan Medical Association 107(4):295-303.

Lindesmith, L., Moe, C., Marionneau, S., Ruvoen, N., Jiang, X., Lindblad, L., Stewart, P., LePendu, J., and Baric, R. 2003. Human susceptibility and resistance to Norwalk virus infection. Nature Medicine 9(5):548-553.

Lindesmith, L., Moe, C., Lependu, J., Frelinger, J.A., Treanor, J., and Baric, R.S. 2005. Cellular and humoral immunity following Snow Mountain virus challenge. Journal of Virology 79:2900-2909.

Lloyd, A.R., Wakefield, D., Boughton, C., and Dwyer, J. 1988. What is myalgic encephalomyelitis? Lancet 1(8597):1286-1287.

Locht, H. and Krogfelt, K.A. 2002. Comparison of rheumatological and gastrointestinal symptoms after infection with *Campylobacter jejuni/coli* and enterotoxigenic *Escherichia coli*. Annals of the Rheumatic Diseases 61:448-452

Loewe, L., Textor, V., and Scherer, S. 2003. High deleterious genomic mutation rate in stationary phase of *Escherichia coli*. Science 302:1558-1559.

Luo, G., Seetharamaiah, G.S., Niesel, D.W., Zhang, H., Peterson, J.W., Prabhakar, B.S., and Klimpel, G.R. 1994. Purification and characterization of Yersinia enterocolitica envelope proteins which induce antibodies that react with human thyrotropin receptor. Journal of Immunology 152(5):2555-2561.

Lyons, C.R., and Wu, T.H. 2007. Animal models of *Francisella tularensis* infection. Annals of the New York Academy of Science 1105:238-265.

Marciano-Cabral, F., and Cabral, G. 2003. *Acanthamoeba* spp. as agents of disease in humans. Clinical Microbiology Reviews 16(2):273-307.

Marciano-Cabral, F., MacLean, R., Mensah, A., and LaPat-Polasko, L. 2003. Identification of *Naegleria fowleri* in domestic water sources by nested PCR. Applied and Environmental *Microbiology* 69(10):5864-5869.

Marciano-Cabral, F., Puffenbarger, R., and Cabral, G.A. 2000.The increasing importance of *Acanthamoeba* infections. Journal of Eukaryotic Microbiology 47(1): 29-36.

Marionneau, S., Ruvoen, N., Le Moullac-Vaidye, B., Clement, M., Cailleau-Thomas, A., Ruiz-Palacois, G., Huang, P., Jiang, X., and Le Pendu, J. 2002. Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals. Gastroenterology 122(7):1967-1977.

Marshall, M.M., Naumovitz, D., Ortega, Y., and Sterling, C.R. 1997. Waterborne protozoan pathogens. Clinical Microbiology Reviews 10(1):67-85.

Maunula, L., Miettinen, I.T., and von Bonsdorff ,C.H. 2005. Norovirus outbreaks from drinking water. Emerging Infectious Diseases 11:1716-1721.

McBride, G.B., Till, D., Ryan, T., Ball, A., Lewis, G., Palmer, S., and Weinstein, P. 2002. Freshwater Microbiology Research Programme. Pathogen Occurrence and Human Health Risk Assessment Analysis. Technical Publication, 93 pp. Wellington, New Zealand: Ministry for the Environment. http://www.mfe.govt.nz/publications/water/freshwater-microbiology-nov02/freshwater-microbiology-nov02.pdf.

McCarthy, N., and Giesecke, J. 2001. Incidence of Guillain-Barre syndrome following infection with *Campylobacter jejuni*. American Journal of Epidemiology 153(6):610-614.

McDaniels, A.E., Wymer, L., Rankin C, and Haugland, R. 2005. Evaluation of quantitative real time PCR for the measurement of *Helicobacter pylori* at low concentrations in drinking water. Water Research 39:4808-4816.

Mead, P.S., Slutsker, L., Dietz,V., McCaig, L.F., Bresee, J.S., Shapiro, C., Griffin, P.M., and Tauxe, R.V. 1999. Food-related illness and death in the United States. Emerging Infectious Diseases 5(5):607-625.

Medema, G.J., Teunis, P.F., Havelaar, A.H., and Haas, C.N. 1996. Assessment of the dose-response relationship of *Campylobacter jejuni*. International Journal of Food Microbiology 30:101-111.

Medema, G., and Smeets, P. 2004. The Interaction Between Quantitative Microbial Risk Assessment and Risk Management in the Water Safety Plan. Kiwa Water Research/Delft University.
http://217.77.141.80/clueadeau/microrisk/uploads/interaction_in_water_safety_plan.pdf.

Meier, P.A., Mathers, W.D., Sutphin, J.E., Folberg, R., Hwang, T., and Wenzel, R.P. 1998. An epidemic of presumed *Acanthamoeba* keratitis that followed regional flooding. Results of a case-control investigation. Archive of Ophthalmology 116:1090-1094.

Mena, K.D., and Gerba, C.P. 2009. Waterborne adenovirus. Reviews of Environmental Contamination and Toxicology 198:133-167.

Messner, M.J., Chappell, C.L., and Okhuysen, P.C. 2001. Water risk assessment for *Cryptosporidium*: a hierarchical Bayesian analysis of human dose response data. Water Research 35(16):3934-3940.

Messner, M.J., Berger, P., Nappier, S.P. 2014. Fractional Poisson - A Simple Dose-Response Model for Human Norovirus. (in press)

Mintz, E.D., Hudson-Wragg, M., Mshar, P., Cartter, M.L., and Hadler, J.L. 1993. Foodborne giardiasis in a corporate office setting. Journal of Infectious Diseases 167(1):250-253.

Moe, C.L., Frelinger, J.A., Heizer, W., and Stewart, P. 2002. Studies of the infectivity of Norwalk and Norwalk-like viruses. Submitted to EPA National Center for Environmental Research (#R826139). Washington, DC: EPA. http://cfpub2.epa.gov/ncer_abstracts/index.cfm/fuseaction/display.abstractDetail/abstract/192/report/F

Moon, H., Chen, J.J., Gaylor, D.W., and Kodell, R.L. 2004. A comparison of microbial dose-response models fitted to human data. Regulatory Toxicology and Pharmacology 40:177-184.

Moon, H., Kim, H-J., Chen, J.J., and Kodell, R.L. 2005. Model averaging using the Kullback Information Criterion in estimating effective doses for microbial infection and illness. *Risk Analysis* 25(5):1147-1159.

Morgan, B. J.T., and Watts, S.A. 1980. On modelling microbial infections. Biometrics 36:317-321.

Morgan, M.G., and Henrion, M. eds. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. New York: Cambridge University Press.

Morgan, U.M., Xiao, L., Sulaiman, I., Weber, R., Lal, A.A., Thomson, R.C., and Deplazes, P. 1999. Which genotypes/species of *Cryptosporidium* are humans susceptible to? Journal of Eukaryotic Microbiology 46(5):42S-43S.

Motarjemi, Y. 2002. Chronic Sequelae of Foodborne Infections. In Foodborne Pathogens, eds. C. d. Blackburn and P. J. McClure. Pp. 501-513. Boca Raton, FL: CRC Press.

Moyer, N.P. 2006. *Aeromonas*. Pp. 81-85 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Moyer, N.P., and Degnan, A.J. 2006. *Shigella*. Pp. 145-148 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Muniesa, M., Jofre, J., García-Aljaro, C., and Blanch, A.R. 2006. Occurrence of *Escherichia coli* O157:H7 and other enterohemorrhagic *Escherichia coli* in the environment. Environmental Science & Technology 40(23):7141-7149.

Mushahwar, I.K. 2008. Hepatitis E virus: molecular virology, clinical features, diagnosis, transmission, epidemiology, and prevention. Journal of Medical Virology 80:646-658.

Nachamkin, I., Allos, B.M., and Ho, T.W. 2000. *Campylobacter jejuni* infection and the association with Guillain–Barre´ syndrome. In *Campylobacter* (2nd edition). Eds. I. Nachamkin and M.J. Blaser. Pp. 155-178. Washington, DC: American Society for Microbiology.

NACMCF (National Advisory Committee on Microbiological Criteria for Foods). 1997. Hazard Analysis and Critical Control Point Principles and Application Guidelines. Adopted August 14, 1997. http://www.fda.gov/Food/FoodSafety/HazardAnalysisCriticalControlPointsHACCP/HACCPPrinciplesApplicationGuidelines/default.htm.

Namata, H., Aerts, M., Faes, C., and Teunis P. 2008. Model averaging in microbial risk assessment using fractional polynomials. Risk Analysis 28(4):891-905.

Nauta, M., Hill, A., Rosenquist, H., Brynestad, S., Fetsch, A., van der Logt, P., Fazil, A., Christensen, B., Katsma, E., Borck, B., and Havelaar, A. 2009. A comparison of risk assessments on *Campylobacter* in broiler meat. International Journal of Food Microbiology 129(2):107-123.

Niyogi, S.K. 2005. Shigellosis. Journal of Microbiology 43(2):133-143.

Noonburg, G.E. 2005. Management of extremity trauma and related infections occurring in the aquatic environment. Journal of the American Academy of Orthopaedic Surgeons 13(4):243-253.

NRC (National Research Council). 1983. Risk Assessment in the Federal Government: Managing the Process. Washington, DC: National Academy Press.

NRC. 2004. Indicators for Waterborne Pathogens. Washington, DC: National Academies Press.

NRC. 2009. Science and Decisions: Advancing Risk Assessment. Washington, DC: National Academies Press.

NZ MFE (Ministry for the Environment, New Zealand). 2003. Microbiological Water Quality Guidelines for Marine and Fresh water Recreational Areas. Available at:http://www.mfe.govt.nz/publications/water/microbiological-quality-jun03/microbiological-quality-jun03.pdf

O'Donoghue, P. 1995. *Cryptosporidium* and cryptosporidiosis in man and animals. International Journal for Parasitology 25:2:139-195.

Okhuysen, P.C., Chappell, C.L., Crabb, J.H., Sterling, C.R., and DuPont, H.L. 1999. Virulence of three distinct *Cryptosporidium parvum* isolates for healthy adults. Journal of Infectious Diseases 180(4):1275-1281.

Okhuysen, P.C., Rich, S.M., Chappell, C.L., Grimes, K.A., Widmer, G., Feng, X., and Tzipori, S. 2002. Infectivity of a *Cryptosporidium parvum* isolate of cervine origin for healthy adults and interferon-gamma knockout mice. Journal of Infectious Diseases 185(9):1320-1325.

Olivieri, A.W., Eisenberg, D., Soller, J., Eisenberg, J., Trussell, R., and Gagliardo, P. 1999. Estimation of pathogen removal in an advanced water treatment facility using Monte Carlo simulation. Water Science and Technology 40:4-5.

OMB (Office of Management and Budget). 2004. Revised Information Quality Bulletin for Peer Review, April 2004. http://www.whitehouse.gov/omb/inforeg/peer_review041404.pdf.

Omenn, G.S., Kessler, A.C., Anderson, N.T., Chiu, P.Y., Doull, J., Goldstein B., Lederberg, J., McGuire, S., Rall, D., and Weldon, V.V. 1997. Framework for Environmental Health Risk Management/Risk Assessment and Risk Management in Regulatory Decision-Making: Final Report (2 vol.). Presidential/Congressional Commission on Risk Assessment and Risk Management. http://www.riskworld.com/Nreports/1997/risk-rpt/pdf/epajan.pdf.

Parkhurst, D.F., Craun, G.F., and Soller, J. 2007. Conceptual Bases for Relating Illness Risk to Indicator Concentrations. In Statistical Framework for Recreational Water Quality Criteria and Monitoring. EPA, Office of Research and Development Monograph.

Parkin, R.T., and Balbus, J.M. 2000. Variations in concepts of "susceptibility" in risk assessment. Risk Analysis 20:603-611.

Parkin, R.T., Soller, J.A., and Olivieri, A.W. 2003. Incorporating susceptible subpopulations in microbial risk assessment: pediatric exposures to enteroviruses in river water. 13: 161-168.

Parsonnet, J., Vandersteen, D., Goates, J., Sibley, R.K., Pritikin, J., and Chang, Y. 1991. Helicobacter pylori infection in intestinal- and diffuse-type gastric adenocarcinomas. Journal of the National Cancer Institute 83(9):640-3.

Petterson, S., Signor, R., Ashbolt, N., and Roser, D. 2006. QMRA Methodology. In Quantitative Microbial Risk Assessment in the Water Safety Plan. In Final Report on the EU MicroRisk Project, Medema, G., Loret, J.-C., Stenstrom, T.A., and Ashbolt, N., eds (European Commission, Brussels), pp. 3-64. Available at: http://www.microrisk.com/uploads/microrisk_qmra_methodology.pdf.

Petterson, S.R., Signor, R.S., and Ashbolt, N.J. 2007. Incorporating method recovery uncertainties in stochastic estimates of raw water protozoan concentrations for QMRA. Journal of Water and Health 5(S1):51-65.

Pinsky, P.F. 2000. Assessment of risks from long term exposure to waterborne pathogens. Environmental and Ecological Statistics 7:155-175.

Plutzer, J., Ongerth, J., and Karanis, P. 2010. *Giardia* taxonomy, phylogeny and epidemiology: facts and open questions. International Journal of Hygiene and Environmental Health 213:321-333.

Pope, J.E., Krizova, A., Garg, A.X., Thiessen-Philbrook, H., and Ouimet, J.A. 2007. *Campylobacter* reactive arthritis: a systematic review. Seminars in Arthritis and Rheumatism 37:48-55.

Pouillot, R., Beaudeau, P., Denis, J.-B., and Derouin, F. for the AFSSA *Cryptosporidium* Study Group. 2004. A Quantitative Risk Assessment of Waterborne Cryptosporidiosis in France Using Second-Order Monte Carlo Simulation. Risk Analysis 24(1):1-17.

Primm, T.P., Lucero, C.A., and Falkinham, J.O. 3rd. 2004. Health impacts of environmental *Mycobacteria*. Clinical Microbiology Reviews 17(1):98-106.

Rambaud, J.C., Halphen, M., Galian, A., and Tsapis, A. 1990. Immunoproliferative small intestinal disease (IPSID): relationships with α-chain disease and "Mediterranean" lymphomas. Springer Seminars in Immunopathology 12(2-3):239-250.

Rangel, J.M., Sparling, P.H., Crowe, C., Griffin, P.M., and Swerdlow, D.L. 2005. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982-2002. Emerging Infectious Diseases 11:603-609.

Reavis, C.J. 2005. Rural health alert: *Helicobacter pylori* in well water. American Academy of Nurse Practitioners 17(7):283-299.

Regli, S., Rose, J.B., Haas, C.N., and Gerba, C.P. 1991. Modeling the risk from *Giardia* and viruses in drinking-water. Journal of the American Water Works Association 83:76-84.

Rendtorff, R.C., 1954a. The experimental transmission of human intestinal protozoan parasites. I. *Endamoeba coli* cysts given in capsules. American Journal of Hygiene 59:196-208.

Rendtorff, R.C., 1954b. The experimental transmission of human intestinal protozoan parasites. II. *Giardia lamblia* cysts given in capsules. American Journal of Hygiene 59:209-220.

Rendtorff, R.C., and Holt, C.J. 1954c. The experimental transmission of human intestinal protozoan parasites. III. Attempts to transmit *Endamoeba coli* and *Giardia lamblia* by flies. The experimental transmission of human intestinal protozoan parasites. I. *Endamoeba coli* cysts given in capsules. American Journal of Hygiene 60:320-326.

Rendtorff, R.C. and Holt, C.J. 1954d. The experimental transmission of human intestinal protozoan parasites. IV. Attempts to transmit *Endamoeba coli* and *Giardia lamblia* by water. American Journal of Hygiene 60:327-338.

Richardson, R.F., Jr., Remler, B.F., Katirji, B., and Murad, M.H. 1998. Guillain-Barre syndrome after *Cyclospora* infection. Muscle Nerve 21(5):669-671.

Rodrigquez-Hernandez, J., Canut-Blasco, A., and Martin-Sanchez, A.M. 1996. Seasonal relevance of *Cryptosporidium* and *Giardia* infections in children attending day care centers in Salamanca (Spain) studied for a period of 15 months. European Journal of Epidemiology 12:291-295.

Rose, J.B., and Gerba, C.P. 1991. Use of risk assessment for development of microbial standards. Water Science and Technology 24:29-34.

Rose, J.B., and Grimes, D.J. 2001. Reevaluation of Microbial Water Quality: Powerful New Tools for Detection and Risk Assessment. Washington, DC: American Academy of Microbiology.

Rose, J.B., Haas, C.N., and Regli, S. 1991. Risk assessment and control of waterborne giardiasis. American Journal of Public Health 81:709-713.

Rose, J.B., and Sobsey, M.D. 1993. Quantitative risk assessment for viral contamination of shellfish and coastal waters. Journal of Food Protection 56(12):1043-1050.

Roseberry, A.M. and Burmaster, D.E. 1992. Lognormal distribution for water intake by children and adults. Risk Analysis 12:99-104.

Roser, D.J., Davies, C.M., Ashbolt, N.J., and Morison, P. 2006. Microbial exposure assessment of an urban recreational lake: a case study of the application of new risk-based guidelines. Water Science and Technology 54**:**245-252.

Roy, S.L., DeLong, S.M., Stenzel, S.A., Shiferaw, B., Roberts, J.M., Khalakdina, A., Marcus, R., Segler, S.D., Shah, D.D., Thomas, S., Vugia, D.J., Zansky, S.M., Dietz, V., and Beach, M.J. 2004. Risk factors for sporadic cryptosporidiosis among immunocompetent persons in the United States from 1999 to 2001. Journal of Clinical Microbiology 42:2944-2951.

Sanchez, J. F., Olmedo, M. C., Pascua, F. J., and Casado, I. 2000. Diabetes insipidus as a manifestation of cerebral toxoplasmosis in an AIDS patient. Revista de Neurologia 30(10):939-940.

Sartwell, P.E. 1950. The distribution of incubation periods of infectious disease. American Journal of Hygiene 51:310-318.

Scallan, E., Hoekstra, R.M., Angulo, F.J., Tauxe, R.V., Widdowson, M.A., Roy, S.L., Jones, J.L., and P.M. Griffin. 2011a. Foodborne illness acquired in the United States — major pathogens. Emerging Infectious Diseases, 17:7-15.

Scallan, E., Griffin, P.M., Angulo, F.J., Tauxe, R.V., and R.M. Hoekstra. 2011b. Foodborne illness acquired in the United States — unspecified agents. Emerging Infectious Diseases, 17:16-22.

Schets, F.M., Schijven, J.F., and de Roda Husman, A.M. 2011. Exposure assessment for swimmers in bathing waters and swimming pools. Water Research 45:2392-2400.

Schiellerup, P., Krogfelt, K.A., and Locht, H. 2008. A comparison of self-reported joint symptoms following infection with different enteric pathogens: Effect of HLA-B27. Journal of Rheumatology 35(3):480-487.

Schiff, G.M., Stefanovic, G.M., Young, E.C., Sander, D.S., Pennekamp, J.K., and Ward, R.L. 1984. Studies of echovirus-12 in volunteers: determination of minimal infectious dose and the effect of previous infection on infectious dose. Journal of Infectious Disease 150(6): 858-866.

Schuster, C.J., Ellis, A.G., Robertson, W.J., Charron, D.F., Aramini, J.J., Marshall, B.J., and Medeiros, D.T. 2005. Infectious disease outbreaks related to drinking water in Canada, 1974-2001. Canadian Journal Public Health 96(4):254-258.

Schuster, M.H., and Visvesvara, G.S. 2004. Pathogenic and opportunistic free-living amoebae: *Acanthamoeba* spp., *Balamuthia mandrillaris*, *Naegleria fowleri*, and *Sappinia diploidea*. FEMS Immunology and Medical Microbiology *5*0(1):1-26.

Schwab, K.J. Astroviruses. 2006. Pp. 259-262 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Schwab, K.J. and Hurst, C. 2006. Human Caliciviruses. Pp. 281-286 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Seto, E.Y., Soller, J.A., and Colford, J.M., Jr. 2007. Strategies to reduce person-to-person transmission during widespread *Escherichia coli* O157:H7 outbreak. Emerging Infectious Diseases 13:860-866.

Sharma S., Sachdeva, P., and Virdi, J.S. 2003. Emerging water-borne pathogens. Applied Microbiology and Biotechnology 61(5-6):424-428.

Shaywitz, S.E., Cohen, P.M., Cohen, D.J., Mikkelson, E., Morowitz, G., and Shaywitz, B.A. 1982. Long-term consequences of Reye syndrome: a sibling-matched, controlled study of neurologic, cognitive, academic, and psychiatric function. Journal of Pediatrics 100(1):41-46.

Sheng, L., Eisenberg, J.N.S., Spiknall, I., Koopman, J.S. 2009. Dynamics and Control of Infections Transmitted from Person to Person through the Environment. American Journal of Epidemiology 170 (2): 257-265.

Shirtliff, M.E., and Mader, J.T. 2002. Acute septic arthritis. Clinical Microbiology Reviews 15(4):527-544.

Sifuentes-Osornio, J., Porras-Cortes, G., Bendall, R.P., Morales-Villarreal, F., Reyes-Teran, G., and Ruiz-Palacios, G.M. 1995. *Cyclospora cayetanensis* infection in patients with and without AIDS: biliary disease as another clinical manifestation. Clinical Infectious Disease 21(5):1092-1097.

Signor, R.S., Roser, D.J., Ashbolt, N.J., and Ball, J.E. 2005. Quantifying the impact of runoff events on microbiological contaminant concentrations entering surface drinking source waters. Journal of Water and Health 3(4):453-68.

Slifko, T.R., Friedman, D., Rose, J.B., and Jankubowski, W. 1997. An in vitro method for detecting infectious *Crytosporidium* oocysts with cell culture. Applied and Environmental Microbiology 63(9):3669-3675.

Slifko, T.R., Huffman, D.E. and Rose, J.B. 1999. A most probable assay for enumeration of infectious *Cryptosporidium parvum* oocycts. Applied and Environmental Microbiology 65(9):3936-3941.

Slifko, T.R., Huffman, D.E., Bertrand, D., Owens, J.H., Jakubowski, W., Haas, C.N. and Rose, J.B. 2002. Comparison of animal infectivity and cell culture systems for evaluation of *Cryptosporidium parvum* oocycts. Experimental Parasitology 101:97-106.

Smeets, P.W., Dullemont, Y.J., Van Gelder, P.H., Van Dijk, J.C., and Medema, G.J. 2008. Improved methods for modelling drinking water treatment in quantitative microbial risk assessment; a case study of *Campylobacter* reduction by filtration and ozonation. Journal of Water and Health 6:301-314.

Smith, J.W., and Wolfe, M.S. 1980. Giardiasis. Annual Review of Medicine 31:373-383.

Smith, A., Reacher, M., Smerdon, W., Adak, G.K., Nichols, G.,and Chalmers, R.M. 2006. Outbreaks of waterborne infectious intestinal disease in England and Wales, 1992-2003. Epidemiology and Infection 134(6):1141-1149.

Snedecor, G.W., and Cochran, W.G. 1989. Statistical Methods, 8th Edition, Iowa State University Press.

Snelling, W.J., Matsuda, M., Moore, J.E., and Dooley, J.S. 2006. Under the microscope: *Arcobacter*. Letters in Applied Microbiology 42:7-14.

Sobsey, M.D. 2006. Hepatitis A Virus. Pp. 273-279 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Soller, J.A., Eisenberg, J.N., and Olivieri, A.W. 1999. Evaluation of Pathogen Risk Assessment Framework. Oakland, CA: Eisenberg, Olivieri and Associates.

Soller, J.A., Olivieri, A., Crook, J., Parkin, R., Spear, R., Tchobanoglous, G., and Eisenberg, J.N.S. 2003. Risk-based approach to evaluate the public health benefit of additional wastewater treatment. Environmental Science & Technology 37:1882-1891.

Soller, J.A., Eisenberg, J., DeGeorge, J., Cooper, R., Tchobanoglous, G., and Olivieri, A. 2006. A public health evaluation of recreational water impairment. Journal of Water and Health 4:1-19.

Soller, J.A., Seto, E.Y., and Olivieri, A.W. 2007. Application of Microbial Risk Assessment Techniques to Estimate Risk Due to Exposure to Reclaimed Waters. WateReuse Foundation, Final Project Report WRF-04-011.

Soller, J.A., and Eisenberg, J.N.S. 2008. An evaluation of parsimony for microbial risk assessment models. Environmetrics 19:61-78.

Soller, J.A. 2009. Potential implications of person-to-person transmission of viral infection to US EPA's Groundwater Rule. Journal of Water and Health 7:208-223.

Soller, J.A., Schoen, M.E., Bartrand, T., Ravenscroft, J., and Ashbolt, N.J. 2010a. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Research: 44(16):4674-91.

Soller, J.A., Bartrand, T., Ashbolt, N.J., Ravenscroft, J., and Wade, T.J. 2010c. Estimating the primary etiologic agents in recreational freshwaters impacted by human sources of faecal contamination. Water Research 44(16): 4736-4747.

Soller, J.A., Schoen, M.E., Varghese, A., Ichida, A.M., Boehm, A.B., Eftim, S., Ashbolt, N.J., Ravenscroft, J.E. 2014. Human Health Risk Implications of Multiple Sources of Faecal Indicator Bacteria in a Recreational Waterbody. Water Research 66:254–264.

Sridharan, S., Mossad, S., and Hoffman, G. 2000.Hepatitis A infection mimicking adult onset Still's disease. Journal of Rheumatology 27(7):1792-1795.

Steinberg, E.B., Greene, K.D., Bopp, C.A., Cameron, D.N., Wells, J.G., and Mintz, E.D. 2001. Cholera in the United States, 1995-2000: trends at the end of the Twentieth Century. Journal of Infectious Diseases 184:799-802.

Stewart, I., Carmichael, W.W., Sadler, R., McGregor, G.B., Reardon, K., Eaglesham, G.K., Wickramasinghe, W.A., Seawright, A.A., and Shaw, G.R. 2009. Occupational and environmental hazard assessments for the isolation, purification and toxicity testing of cyanobacterial toxins. Environmental Health 8:52.

Stone, P. 2006. EU Private International Law: Harmonization of Laws (Elgar European Law Series). Cheltenham, UK: Edwin Elgar Publishing Limited.

Sur, D., Ramamurthy, T., Deen, J., and Bhattacharya, S.K. 2004 Shigellosis: challenges and management issues. Indian Journal of Medical Research 120:454-462.

Tanaka, H., Asano, T., Schroeder, E.D., and Tchobanoglous, G. 1998. Estimating the safety of wastewater reclamation and reuse using enteric virus monitoring data. Water Environment Research 70(1):39-51.

Teunis, P.F., van der Heijden, O.G., van der Giessen, J.W.B., and Havelaar, A.H. 1996. The Dose-Response Relation in Human Volunteers for Gastro-intestinal Pathogens. Report No. 284550002. Bilthoven, The Netherlands: RIVM (National Institute of Public Health and the Environment).

Teunis, P.F.M., and Havelaar, A.H. 1999. *Cryptosporidium* in Drinking Water: Evaluation of the ILSI/RSI Quantitative Risk Assessment Framework. Report No. 284 550 006. Bilthoven, The Netherlands: RIVM.

Teunis, P.F.M., and Havelaar, A. 2000. The beta-Poisson model is not a single hit model. Risk Analysis 20(4):513-520.

Teunis, P.F., Chappell, C.L., and Okhuysen, P.C. 2002 *Cryptosporidium* dose response studies: Variation between isolates. Risk Analysis 22(1):175-183.

Teunis, P., Takumi, K., and Shinagawa, K. 2004. Dose response for infection by *Escherichia coli* O157:H7 from outbreak data. Risk Analysis 24(2):401-407.

Teunis, P.F.M., van den Brandhof, W., Nauta, M., Wagenaar, J., van den Kerkhof, H., and Van Pelt, W. 2005. A reconsideration of the *Campylobacter* dose-response relation. Epidemiology and Infection 133:583-592.

Teunis, P.F.M., Ogden, I.D., and Strachan, N.J.C. 2008a. Hierarchical dose response of *E. coli* O157:H7 from human outbreaks incorporating heterogeneity in exposure. Epidemiology and Infection 136(6):761-770.

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

Teunis, P.F.M. 2009. Uncertainty in Dose Response from the Perspective of Microbial Dose. Ch. 6 in Uncertainty Modeling in Dose Response. Hoboken, NJ: Wiley.

Teunis, P.F., Kasuga, F., Fazil, A., Ogden, I.D., Rotariu, O., and Strachan, N.J. 2010. Dose-response modeling of Salmonella using outbreak data. International Journal of Food Microbiology 144(2): 243-249.

Thebpatiphat, N., Hammersmith, K.M., Rocha, F.N., Rapuano, C.J., Ayres, B.D., Laibson, P.R., Eagle, R.C., Jr., and Cohen, E.J. 2007. *Acanthamoeba* keratitis: a parasite on the rise. Cornea 26(6):701-706.

Thoeye, C., Van Eyck, K., Bixio, D., Weemaes, M., and De Gueldre, G. 2003. Methods Used for Health Risk Assessment in State of the Art Report: Health Risks in Aquifer Recharge Using Reclaimed Water. Geneva, Switzerland: WHO. http://www.who.int/water_sanitation_health/wastewater/en/wsh0308chap4.pdf.

Thomson, G.T.D., Derubeis, D.A., Hodge, M.A., Rajanayagam, C., and Inman, R.D. 1995. Post-*Salmonella* reactive arthritis: late clinical sequelae in a point source cohort. American Journal of Medicine 98(1):13-21.

Thompson R.C. 2000. Giardiasis as a re-emerging infectious disease and its zoonotic potential. International Journal of Parasitology 30:1259-1267.

Toranzos, G.A., et al. (2006). *Vibrio cholerae*. Pp. 153-156 in Waterborne Pathogens. AWWA Manual M48, 2nd Edition. Denver, CO: American Water Works Association.

Trachoo, N. 2003. *Campylobacter jejuni*: an emerging pathogen. Songklanakarin Journal of Science and Technology 25(1):141-157.

Tuncay, S., Delibaş, S., Inceboz, T., Over, L., Oral, A.M., Akisü, C., and Aksoy, U. 2008. An outbreak of gastroenteritis associated with intestinal parasites. Türkiye parazitolojii dergisi 32(3):249-252.

Tupchong, M., Simor, A., and Dewar, C. 1999. Beaver fever—a rare cause of reactive arthritis. Journal of Rheumatology 26:2701-2702.

USDA (U.S. Department of Agriculture). 1998. *Salmonella Enteritidis* Risk Assessment Shell Eggs and Egg Products. http://www.fsis.usda.gov/ophs/risk/pdfrisk1.pdf.

U.S. Census Bureau. 2011a. Table 11. Resident Population by Race, Hispanic Origin, and Single Years of Age: 2009. U.S. Census Bureau, Statistical Abstract of the United States: 2011. Available at: http://www.census.gov/compendia/statab/2011/tables/11s0010.pdf. Accessed on October 18, 2012.

U.S. Census Bureau. 2011b. Table 1248. Participation in selected sports activities: 2008. U.S. Census Bureau, Statistical Abstract of the United States: 2011. Available at: http://www.census.gov/compendia/statab/2011/tables/11s1248.pdf. Accessed on October 18, 2012.

U.S. EPA. 1989. Risk Assessment Guidance for Superfund. EPA-540/1-89/002. Washington, DC. http://www.epa.gov/oswer/riskassessment/ragsa/index.htm.

U.S. EPA. 1992. Framework for Ecological Risk Assessment. EPA/630/R-92/001. Washington, DC

U.S. EPA. 1995a. Guidance for Risk Characterization. U.S. Environmental Protection Agency, Science Policy Council. Washington, DC. http://www.epa.gov/OSA/spc/pdfs/rcguide.pdf

U.S. EPA. 1995b. The EPA's Environmental Justice Strategy. http://www.epa.gov/compliance/resources/policies/ej/ej_strategy_1995.pdf.

U.S. EPA.1997a. Exposure Factors Handbook. EPA/600/P-95/002Fa. Office of Research and Development, National Center for Environmental Assessment. Washington, DC.

U.S. EPA. 1997b. Guiding Principles for Monte Carlo Analysis. EPA/630/R-97/001.

U.S. EPA. 1998a. *Giardia*: Human Health Criteria Document. EPA-823-R-002.

U.S. EPA. 1998b. Guidelines for Ecological Risk Assessment. May 14, 1998, Federal Register 63(93):26846-26924. EPA/630/R-95/002.F.

U.S. EPA. 1999b. *Legionella:* Human Health Criteria Document. EPA-822-R-99-001. Washington, DC.

U.S. EPA. 1999c. Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments. EPA/630/R-98/004. Risk Assessment Forum, Washington, DC.

U.S. EPA. 2000a. EPA Science Policy Council Peer Review Handbook. EPA-100-B-00-001. Washington, DC.

U.S. EPA. 2000b. Science Policy Council Risk Characterization Handbook. EPA-100-B-00-002. Washington, DC. http://www.epa.gov/osa/spc/pdfs/rchandbk.pdf.

U.S. EPA. 2000c. Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health. EPA-822-B-00-004. http://www.epa.gov/waterscience/humanhealth/method/complete.pdf.

U.S. EPA. 2000d. Report to Congress EPA Studies on Sensitive Subpopulations and Drinking Water Contaminants. EPA 815-R-00-015. http://www.epa.gov//safewater/standard/rtc_sensubpops.pdf.

U.S. EPA. 2000e. Guidelines for Preparing Economic Analyses. National Center for Environmental Economics, EPA/240/R-00/003. Washington, DC.

U.S. EPA. 2001. Protocol for Developing Pathogen TMDLs. EPA 841-R-00-002. http://www.epa.gov/owow/tmdl/pathogen_all.pdf.

U.S. EPA. 2002a. Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity, of Information Disseminated by the Environmental Protection Agency. EPA/260R-02-008. http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf.

U.S. EPA. 2002b. Lessons Learned on Planning and Scoping for Environmental Risk Assessments, EPA Science Policy Council. Washington, DC. http://epa.gov/osp/spc/handbook.pdf.

U.S. EPA. 2002c. *Mi*crobiological Risk Assessment Framework Workshop Tools, Methods, and Approaches. Prepared by ICF Consulting, Inc.

U.S. EPA. 2002d. Child-Specific Exposure Factors Handbook. EPA-600-P-00-002B.

U.S. EPA. 2002e. National Primary Drinking Water Regulations: Long Term 1 Enhanced Surface Water Treatment Rule. 40CFR Parts 9, 141 and 142, and *Federal Register* 67(9), January 14, 2002. http://www.epa.gov/safewater/mdbp/lt1eswtr.html.

U.S. EPA. 2003a. National Primary Drinking Water Regulations: Long Term 2 Enhanced Surface Water Treatment Rule Proposal. Federal Register 68:154.

U.S. EPA. 2003b. Economic Analysis for the Long Term 2 Enhanced Surface Water Treatment Rule Proposal.

U.S. EPA. 2003c. Microbiological Risk Assessment Framework: Problem Formulation Workshop. Prepared by ICF Consulting, Inc.

U.S. EPA. 2003d. Movement and Longevity of Viruses in the Subsurface. National Risk Management Research Laboratory. EPA/540/S-03/500. http://www.epa.gov/ada/download/issue/540S03500.pdf.

U.S. EPA. 2003e. Integrated Risk Information, Glossary of IRIS Terms. http://www.epa.gov/iris/gloss8.htm.

U.S. EPA. 2003f. Framework for Cumulative Risk Assessment. EPA/630/P-02/001F. Office of Research and Development, National Center for Environmental Assessment, Washington, DC. http://www.epa.gov/raf/publications/pdfs/frmwrk_cum_risk_assmnt.pdf.

U.S. EPA. 2003g. Public Involvement Policy of the U.S. Environmental Protection Agency. EPA 233-B-03-002. http://www.epa.gov/publicinvolvement/pdf/policy2003.pdf.

U.S. EPA. 2003h. Health Effects Support Document for *Acanthamoeba*. EPA-822-R-03-012. Washington, DC. http://www.epa.gov/safewater/ccl/pdfs/reg_determine1/support_cc1_acanthamoeba_healtheffcts.pdf.

U.S. EPA. 2004b. Air Toxics Risk Assessment Reference Library: Volume 1 Technical Resource Manual. EPA-453-K-04-001A. http://www.epa.gov/ttn/fera/risk_atra_main.html.

U.S. EPA. 2004c. Developing Dynamic Infection Transmission Models for Microbial Risk Assessment Applications. EPA-NCEA-C-1463.

U.S. EPA. 2004d. Risk Assessment Principles and Practices. EPA/100/B-04/001. Office of the Science Advisor Staff Paper. http://www.epa.gov/OSA/pdfs/ratf-final.pdf.

U.S. EPA. 2005a. Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001B. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=116283.

U.S. EPA. 2005b. Guidance on Selecting Age Groups for Monitoring and Assessing Childhood Exposures to Environmental Contaminants. EPA/630/P-03/003F. National Center for Environmental Assessment, Washington, DC.

U.S. EPA. 2006a. National Primary Drinking Water Regulations: Long Term 2 Enhanced Surface Water Treatment Rule – Final. Federal Register 71(3). http://www.epa.gov/safewater/disinfection/lt2/regulations.html.

U.S. EPA. 2006b. National Primary Drinking Water Regulations: Ground Water Rule. 40CFR Parts 9, 141 and 142, Federal Register 71(216). http://www.epa.gov/safewater/disinfection/gwr/.

EPA. 2006c. Occurrence and Monitoring Document for Final Ground Water Rule. USEPA 815-R-06-012. Office of Water, Washington DC.

U.S. EPA. 2006d. *Aeromonas:* Human Health Criteria Document. Washington, DC: Office of Water.

U.S. EPA. 2006e. EPA Science Policy Council Peer Review Handbook. EPA/100/B-06/002. Washington, DC.

U.S. EPA. 2007a. Thesaurus of Terms Used in Microbiological Risk Assessment. EPA Office of Water. http://www.epa.gov/waterscience/criteria/humanhealth/microbial/thesaurus/index.html.

U.S. EPA. 2007b. Compendium of Prior and Current Microbial Risk Assessment Methods for Use as a Basis for the Selection, Development, and Testing of a Preliminary Microbial Risk Assessment Framework. EPA/600/R-07/129. National Homeland Security Research Center. http://www.epa.gov/NHSRC/pubs/600r07129.pdf.

U.S. EPA. 2008. Child-Specific Exposure Factors Handbook. Final Report. EPA/600/R-06/096F. U.S. Environmental Protection Agency, Washington, DC.

U.S. EPA 2009a. Review of Zoonotic Pathogens in Ambient Waters. EPA 822-R-09-002. http://water.epa.gov/scitech/swguidance/waterquality/standards/criteria/health/recreation/upload/2009_07_16_criteria_recreation_zoonoticpathogensreview.pdf.

U.S. EPA. 2009b. Using Probabilistic Methods to Enhance the Role of Risk Analysis in Decision Making With Case Study Examples (External Review Draft). EPA/100/R-09/001. Risk Assessment Forum, Washington, DC.

U.S. EPA. 2010. Quantitative Mircobial Risk Assessment to Estimate Illness in Freshwater Impacted by Agricultural Animal Sources of Fecal Contamination, EPA 822-R-10-005, Office of Science and Technology, Washington, D.C.

U.S. EPA. 2011. Exposure Factors Handbook: 2011 Edition, EPA/600/R-090/052F. http://www.epa.gov/ncea/efh/pdfs/efh-complete.pdf.

U.S. EPA. 2012a. Framework for Human Health Risk Assessment to Inform Decision Making. EPA External Review Draft. 601-D12-001. http://www.epa.gov/raf/files/framework-document-7-13-12.pdf.

U.S. EPA. 2012b. Peer Review Handbook, 3rd edition. EPA/100/B-06/002. http://www.epa.gov/peerreview/pdfs/peer_review_handbook_2012.pdf.

U.S. EPA. 2012c. Recreational Water Quality Criteria. 820-F-12-058. http://water.epa.gov/scitech/swguidance/standards/criteria/health/recreation/index.cfm

U.S. EPA/USDA. 2012. Microbial Risk Assessment Guideline Pathogenic Microorganisms with Focus on Food and Water. EPA/100/J-12/001 and USDA/FSIS/2012-001. http://www.epa.gov/raf/files/mra-guideline-july-final.pdf.
U.S. EPA. 2014. Overview of Technical Support Materials: A Guide to the Site-Specific Alternative Criteria TSM documents. EPA-820-R-14-010.

Visvesvara, G.S., and Moura, H. 2006. *Acanthamoeba* spp. In Waterborne Pathogens. Pp. 141-144. Denver, CO: American Water Works Association.

Visvesvara, G. S., Moura, H., and Schuster, F. L. (2007). Pathogenic and opportunistic free-living amoebae: *Acanthamoeba* spp., *Balamuthia mandrillaris*, *Naegleria fowleri*, and *Sappinia diploidea*. FEMS Immunology and Medical Microbiology 50(1):1-26.

Vivier, J.C., Ehlers, M.M., and Grabow, W.O. 2004. Detection of enteroviruses in treated drinking water. Water Reearch 38:2699-2705.

Wang, C., Yuan, Y., and Hunt, R.H. 2007. The association between *Helicobacter pylori* infection and early gastric cancer: a meta-analysis. American Journal of Gastroenterology 102(8):1789-1798.

Walter, J.E., and Mitchell, D.K. 2000. Role of astroviruses in childhood diarrhea. Current Opinion in Pediatrics 12(3):275-279.

Weir, E. 2000. *Escherichia coli* O157:H7. Canadian Medical Association Journal 163(2):205.

Weir, M.H., Pepe Razzolini, M.T., Rose, J.B., and Masago, Y. 2011. Water reclamation redesign for reducing *Cryptosporidium* risks at a recreational spray park using stochastic models. Water Research 45(19): 6505-6514.

Westrell, T., Bergstedt, O., Stenström, T.A., and Ashbolt, N.J. 2003. A theoretical approach to assess microbial risks due to failures in drinking water systems. International Journal of Environmental Health Research 13:181-197.

Westrell, T. 2004. Microbial Risk Assessment and its Implications for Risk Management in Urban Water Systems. Doctoral Thesis, Linköping University, Sweden: the Tema Institute, Department of Water and Environmental Studies. http://www.diva-portal.org/liu/abstract.xsql?dbid=4880.

WHO (World Health Organization). 2001. Water Quality: Guidelines, Standards and Health: Assessment of Risk and Risk Management for Water-Related Infectious Disease. Eds. L. Fewtrell, J. Bartram. Published on behalf of IWA Publishing, WHO and Swedish Institute for Infectious Disease Control. http://www.who.int/water_sanitation_health/dwq/whoiwa/en/index.html.

WHO. 2004. Waterborne Zoonoses: Identification, Causes and Control. World Health Organization (WHO); Cotruvo, J.A., Dufour, A., Rees, G., Bartram, J., Carr, R., Cliver, D.O., Craun, G.F., Fayer, R., Gannon, V.P.J (eds). IWA Publishing: London, UK. ISBN: 1 84339 058 2. http://www.who.int/water_sanitation_health/diseases/zoonoses/en/.

WHO. 2009. Hepatitis E. Fact Sheet #280 (revised January 2005) http://www.who.int/mediacentre/factsheets/fs280/en/index.html.

Williams, T. 1965. The basic birth-death model for microbial infections. Journal of the Royal Statistical Society Part B 27(2):338-360.

Willner, I.R., Uhl, M.D., Howard, S.C., Williams, E.Q., Riely, C.A., and Waters, B. 1998. Serious hepatitis A: an analysis of patients hospitalized during an urban epidemic in the United States. Annals of Internal Medicine (2):111-114.

Wolf, M.W., Misaki, T., Bech, K., Tvede, M., Silva, J. E., and Ingbar, S.H. 1991. Immunoglobulins of patients recovering from *Yersinia enterocolitica* infections exhibit Graves' disease-like activity in human thyroid membranes. Thyroid 1(4):315-320.

Wolfe, M.S. 1990. Clinical Symptoms and Diagnosis by Traditional Methods. In Human Parasitic Diseases, Vol. 3, Giardiasis, ed. E.A. Meyer, pp. 175-186, Amsterdam: Elsevier Science Publishers.

Wong, T.Y., Tsui, H.Y., So, M.K., Lai, J.Y., Lai, S.T., Tse, C.W., and Ng T.K. 2000. *Plesiomonas shigelloides* infection in Hong Kong: retrospective study of 167 laboratory-confirmed cases. Hong Kong Medical Journal 6(4):375-380.

Wooldridge, M., and Schaffner, D., eds. 2008. Qualitative Risk Assessment. In Microbial Risk Analysis of Foods. Washington, DC: ASM Press.

Xiao, L., Morgan, U.M., Fayer, R., Thompson, C., and Lal, A.A. 2000. *Cryptosporidium* systematics and implications for public health. Parisitology Today 16(7):287-292.

Yoder, J., Roberts, V., Craun, G. F., Hill, V., Hicks, L. A., Alexander, N. T., Radke, V., Calderon, R.L., Hlavsa, M.C., Beach, M.J., and Roy, S.L. 2008. Surveillance for waterborne disease and outbreaks associated with drinking water and water not intended for drinking--United States, 2005-2006. MMWR Surveillance Summaries 57(9):39-62.

Yu, S.Z. 1989. Drinking Water and Primary Liver Cancer. Pp. 30-37 in Primary Liver Cancer, eds. Y. Tang, M.C. Wu, S.S. Xia, New York: China Academic Publishers.

Yu, S.Z. 1995. Primary prevention of hepatocellular carcinoma. Journal of Gastroenterology and Hepatology 10(6):674-682.

Yu, S., Zhao, N., and Zi, X. 2001. The relationship between cyanotoxin (microcystin, MC) in pond-ditch water and primary liver cancer in China. Zhonghua Zhong Liu Za Zhi 23(2): 96-99.

Yuki, N. 2001. Infectious origins of, and molecular mimicry in Guillain-Barré and Fisher syndromes. The Lancet Infectious Diseases 1(1):29-37.

Zelner J., King, A.A., Moe C.L., and J.N.S. Eisenberg. 2010. How Infections Propagate After Point Source Outbreaks: An Analysis of Secondary Norovirus Transmission. Epidemiology 21(5).

Appendix A. Possible Future MRA Goals and Research Needs

Some examples of possible long-term development goals for microbial risk assessment include the following:

A.1. Exposure Assessment

- Considering how **animal reservoirs of disease** might be incorporated into MRA.
- Developing better methods to account for the **heterogeneous distribution of microorganisms and the potential fluctuations in density** of microorganisms in the environment (spatial heterogeneity and temporal fluctuations).
- Developing methods to address **relative source contribution for microbial risks**; that is, evaluating the relative contribution of drinking water and other pathways (such as food, swimming/recreational, and other environmental exposures) to the total disease risk from all sources. This can also include the development of microbial bioaccumulation factors for organisms that can accumulate human pathogens and are eaten by humans raw or partially cooked (e.g., shellfish). This can be based on an understanding of "disease ecology" (e.g., consider all exposures that result in a given health endpoint) rather than on common assumptions that tend to simplify that understanding (e.g., exposure via a single pathway).

A.2. Human Health Effects Assessment and Dose-Response

- Developing dose-response models that consider situations where populations may be repeatedly exposed to certain microbial pathogens over time (**discrete versus continuous dose and exposure**). These models may include susceptibility and immunity variation and life stages.
- Developing **criteria for the use of animal model results** for derivation of dose-response models. Improved methods to extrapolate animal dose-response information to human dose-response models should be pursued, as well as better ways to address the uncertainty involved in such extrapolations (such as differences in health effects between humans and animals).
- Exploring the issue of whether **threshold or nonthreshold dose-response models** are most appropriate for various pathogen-host combinations.
- Developing **biologically-based mechanistic models** (such models are being developed but are not yet available).
- Developing methods to investigate **dose-response relationships for immunocompromised and other more sensitive populations**. This can include outbreak related studies, epidemiological studies, or studies with immunocompromised animal models.

A.3. Risk Estimation and Risk Characterization

- Developing methods to address **cumulative risks from exposure to multiple pathogens** and to pathogens and certain chemicals.
- Developing additional methods for considering possible **lifetime, cumulative risk** from exposure to one or more pathogens.
- Developing a framework for conducting **community-based** cumulative risk assessment for microbial hazards.
- Developing methods for estimating risks of **chronic or secondary sequelae**.
- Developing methods for comparing risks among different pathogens and different exposures (**comparative risk**). Common metrics that provide a basis for such comparisons (e.g., to compare *Vibrio vulnificus* and *E. coli* O157:H7) should be explored.
    - The use of disability-adjusted life years (DALYs), which measure and compare the effects of disease burden on a population, use morbidity (years lived with a disability), mortality (years of lost life), and standardized life expectancy to calculate a DALY value for a given disease for a defined population, is one method for comparing risks (Gold et al., 2002).
    - Quality-adjusted life years (QALYs) are a method for assigning a numerical value for quality of life and translating that numerical value to a monetary measure (WHO, 2001).
- Conducting research on the appropriate use of **adjustment factors for microbial risk assessment**. The circumstances for using such factors and the criteria to determine the magnitude of the factors and where they can be applied should be considered.
- Developing additional **model validation methods** to compare the results of the risk assessment with "reality." If few data exist for this comparison, after the risk assessment is conducted endpoints should be monitored so the model can be validated in the future.
- Further developing **qualitative assessment methods**, because quantitative data are not always available.
- Improving the application of **risk assessment as a predictive tool in developing prevention strategies**.
- Further developing methods and models for incorporating information on **secondary transmission**.
- Further developing methods and models for **incorporating information on immune status**. For some dose-response datasets where infection is the endpoint, this may be difficult because immunity affects illness rather than infection (e.g., as observed for *Giardia*).

A.4. General research needs to improve MRA include the following:

- more information on mechanisms of infection and virulence factors;
- data on variation among different hosts and pathogens;
- data on the effect of environment on pathogen growth, survival, and death;
- data from longer time frames in order to account for longer-term weather cycles (e.g., el Niño);
- data on changing land use patterns advancement

- improved sampling, detection, quantification methods, and viability/infectivity assays; and
- continued development of a thesaurus or lexicon of risk assessment terms to facilitate the evolution of terminology.

Appendix B. Other Risk Frameworks that are Consistent with the MRA Tools Framework

- The WHO *State of the Art Report: Health Risks in Aquifer Recharge Using Reclaimed Water* has a chapter on methods for health risk assessment that includes a process diagram for risk assessment (Thoeye et al., 2003).
- Quantitative microbial risk assessment (QMRA) approaches, along with summaries of six research papers related to health risks from infectious microorganisms transmitted via urban water and wastewater systems, are presented in this dissertation (Westrell, 2004). Discussions of susceptibility and immunity, sensitive subpopulations, secondary transmission, dynamic modeling, and health indices are also included.
- Rose and Grimes (2001) present a flow diagram for conducting a screening level risk assessment (preliminary risk assessment) that advances users through nine questions to ask during the planning of a screening level risk assessment. Molecular tools for characterizing and identifying microorganisms are also reviewed.
- Medema and Smeets (2004) discuss the interaction between QMRA and the risk management aspects of the WHO Water Safety Plan.
- The Canadian report, *Microbial Risk Assessment as a Foundation for Informed Decision-Making* (Fazil et al., 2005), presents MRA in its larger context by discussing enabling legislation, policy scrutiny, and international trade agreements and standards. The "current status" as well as "the way ahead" is presented for prioritization and coordination; methods and tool development; guidance documents (qualitative, technical, and methodology); training for risk assessors; and risk-based decision-making, peer review, and integration of risk communication.
- The aim of the MICRORISK Project (www.microrisk.com) is to develop a MRA process that contributes to the decision-making process for risk management of drinking water. The elements of the framework are the Quantitative Microbial Risk Assessment (QMRA) and Hazard Analysis and Critical Control Points (HACCP). Funding entities include the following: collaborative water utilities in the Netherlands (BTO), U.K. Water Industry Research, and the Australian Commonwealth Government Department of Education Science and Technology.
- The Center for Advancing Microbial Risk Assessment (CAMRA) is the Homeland Security Center of Excellence jointly established with the U.S. EPA to develop scientific knowledge on the fate and risk of potential bioterrorist and other high priority infectious agents. (http://www.camra.msu.edu/).
- FDA hosts iRisk which is a web-based system designed to analyze data concerning microbial and chemical hazards in food and return an estimate of the resulting health burden on a population level.(https://irisk.foodrisk.org/)

The Interagency *MRA Guideline* was released after major development of this MRA Tools document. These two documents are compatible. In addition, the external review draft of this MRA Tools document was used in the development of the Interagency *MRA Guideline* (U.S. EPA/USDA, 2012).

Appendix C. Dose-Response Models

C.1. Alternate Dose-Response Models

C.1.1. Empirical Models

To date, the majority of studies in dose-response modeling and MRA for waterborne pathogens have employed the exponential and beta-Poisson dose-response models. These models are mechanistic (based on models of biologically-plausible processes), relatively simple, and have provided good fits to many of the data sets for which they have been applied. Other models have also been proposed and used as components of MRAs, particularly in the assessment of risks associated with food. These alternative models are empirical (i.e., not derived based on consideration of biological processes) and as such, their validity outside the data range for which their parameters are estimated is unknown, and extrapolation with these models is not recommended (Buchanan et al., 2000). In dose-response model selection, preference is given to biologically plausible, mechanistic models such as the exponential and beta-Poisson over empirical models. Other researchers (e.g., Coleman and Marks, 1998) have suggested that the exponential and beta-Poisson models are not substantially different from empirical models. However, this document does not adopt that position because exponential and beta-Poisson models may be derived from basic considerations of the infection process, because the models may be adapted to include other processes within the infection process, and because of the demonstrated success in fitting the exponential and beta-Poisson models to available data.

Buchanan et al. (2000), Moon et al. (2004), Holcomb et al. (1999), and Haas et al. (1999) compared the exponential and beta-Poisson models to various empirical dose-response models proposed for use in dose-response modeling for foodborne pathogens. The empirical models assessed in those studies are summarized in Table C-1. Presentation of these models in this document is intended to provide a thorough description of models that have been used. In their comparison of models, Moon et al. (2004) found that three-parameter models did not yield significant improvements in fit over two-parameter models and that among two-parameter models, predictions in the low-dose range were markedly different between models. Buchanan et al. (2000) suggest the development of mechanistic models with consideration of factors important in the infection process as a route to more accurate dose-response models that may be extrapolated outside the range of available data. Holcomb et al. (1999) noted that among the empirical and mechanistic dose-response models compared, only the three-parameter Weibull-gamma model provided goodness of fit for data sets of dose-response data for four different pathogens and that different models predicted very different low-dose-response. These observations lead the authors to suggest that continued dose-response model development and evaluation is necessary.

Table C-1. Empirical Dose-Response Models

Model	Equation	Parameters
Weibull-Gamma	$P(d) = 1 - \left(1 + d^b / \beta\right)^{-\alpha}$	Three parameter model: *b*, α, β.
Weibull	$P(d) = 1 - \exp(-ad^b)$	Two parameter model: *a*, *b*
Gompertz[1]	$P(d) = 1 - \exp\left[-\exp\left(a + b\, f(d)\right)\right]$	Two parameter model: *a*, *b* *f*(*d*) denotes a transformation (e.g., log)
Log-normal[2]	$P(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\ln d - \alpha)/\beta} \exp\left(-\frac{1}{2}t^2\right) dt$	Two parameter model: α, β
Log-logistic	$P(d) = 1/\left\{1 + \exp\left[-(\ln d - \alpha)/\beta\right]\right\}$	Two parameter model: α, β
Exponential -Gamma	$P(d) = 1 - \exp(-\gamma d)/\left(1 + d^b / \beta\right)^{\alpha}$	Three parameter model: α, β, γ
Weibull-exponential	$P(d) = 1 - \exp\left(-\alpha d^{\gamma}\right)/\left(1 + d^{\gamma} / \beta\right)$	Three parameter model: α, β, γ
Shifted Weibull	$P(d) = \begin{cases} 1 - \exp\left\{-\left[(d-\alpha)/\beta\right]^{\gamma}\right\} & d \geq \alpha \\ 0 & 0 \leq d < \alpha \end{cases}$	Three parameter model: α, β, γ

[1] When the function *f* is the natural log transformation of dose, this model is referred to as an "extreme value" model in Moon et al. (2004).
[2] Referred to as the log-probit model in Haas et al. (1999).

Another model, not included in Table C-1 because its origin and functional form differ substantially from those models in the table, was proposed by Brynestad et al. (2008) and subsequently applied by Nauta et al. (2009). The model is termed a "sigmoidal model" and is suggested as an alternative to the beta-Poisson or the hazard function model proposed by Teunis et al. (1999) for the predicted incidence of illness given infection. Nauta et al. (2009) suggest that the beta-Poisson and exponential models are valid for predicting rates of infection but not for rates of illness. This assertion notwithstanding, the exponential and beta-Poisson models have been used in many cases for development of dose-response models in which illness is an endpoint. Rather than using a biological basis for model selection, Brynestead et al. (2008) selected the sigmoidal model based on its simplicity and ability to incorporate expert judgment (specifics of which are not provided herein) and suggest that these bases make the model an improvement over other dose-response models for predicting response at low dose. Their dose-response model (probability of illness for a given dose) is given by:

$$p(d) = (\max - \min)\left(1 + 10^{\log(\mathit{IllD}_{50}) - d}\right) HS \qquad [C\text{-}1]$$

where *HS* is the Hill slope, given as:

$$HS = \frac{\log\left[q/(100-q)\right]}{\log\left(IllD_q / IllD_{50}\right)} \quad \text{[C-2]}$$

q is the chance of becoming ill (selected as 1%), $IllD_{50}$ is the median infectious dose at which 50% of the exposed population becomes ill, $IllD_q$ is the dose at which *q*% of the exposed population is expected to become ill, and log refers to $\log_{10}$. The functional form of the sigmoidal relation is such that the probability of illness rises sharply from a very small value at the dose expected to produce illness in *q%* of the exposed population (1% in the model of Brynestead et al., 2008). Based on published data, Brynestead et al. (2008) estimated that $IllD_{50}$ and $IllD_1$ were uniformly distribution in the ranges 500 to 800 and 2000 to 6000 organisms, respectively. Risk estimates for *Campylobacter* infection related to food preparation appear high for both the Hill slope model and an alternative dose-response model, with the Hill slope model providing a lower, but unrealistic estimate of the number of illnesses. This finding could indicate that current dose-response models over-predict the incidence of illness or that exposure models over predict the incidence and ingestion of *Campylobacter* in the food chain in Germany. Alternatively, the findings could be an artifact of the high uncertainty inherent to the epidemiological data to which QMRA model results were compared.

In their review of foodborne *Campylobacter* illness QMRAs, Nauta et al. (2009) compared alternative *Campylobacter* dose-response models (illness endpoint) and observed that the sigmoidal model predicts much lower *illness* probability at low dose than alternative published models. This observation is consistent with the chosen form of the model. Based on the assumptions of Brynestead et al. (2008), the sigmoidal illness incidence model reported by Nauta et al. (2009) used a 1% illness incidence as the lower end of the illness dose-response relation, arriving at the following dose-response model:

$$p_{ill}(d) = \frac{1}{\left(\frac{IllD_{50}}{d}\right)^a + 1} \quad \text{[C-3]}$$

where $IllD_{50}$ is the dose at which 50% of the exposed population becomes ill, *d* is the ingested dose and *a* is given by

$$a = \frac{\left[\frac{\ln(0.99)}{\ln(0.01)}\right]}{\left[\frac{\ln(IllD_{50})}{\ln(IllD_1)}\right]} \quad \text{[C-4]}$$

The comparison of QMRAs of *Campylobacter* by Nauta et al. (2009) allowed the authors to conclude that the *Campylobacter* dose-response model remains unknown, particularly given potential variations in the ability of different strains of *Campylobacter* to initiate infection or illness, influences that food matrix or other environmental factors may exert on the incidence of illness, and the difference in response for different subpopulations.

Selection of dose-response models requires comparison of fits of the models to data and comparison of fits of more highly-parameterized models with those of models with fewer parameters. When maximum likelihood estimation is used for determining the model parameters, fits are compared on the basis of the deviances at the parameter values providing the best fit of the model to the data. In general, models with more parameters are selected over models with fewer parameters only when the improvement in fit of the model with more parameters over that of the model with fewer parameters is statistically significant.

C.1.2. Threshold Models

Threshold models (i.e., models that assume more than one organism is required to initiate infection) can be derived under slightly different assumptions than those used to develop the exponential and beta-Poisson dose-response models. Assuming pathogens in an ingested dose are drawn from a homogeneous distribution (Poisson distribution) and each pathogen has an equal, independent probability that it can initiate an infectious focus, the probability of infection by $k_{\min}$ organisms is (Haas et al., 1999):

$$P(\text{infection} \mid d) = \Gamma(k_{\min}, r\,d) \qquad \text{[C-5]}$$

where r is a parameter of the distribution and Γ denotes the gamma cumulative probability distribution function. This simple threshold model is a two parameter model whose parameters may be determined via standard statistical techniques such as maximum likelihood estimation (MLE). Deterministic models have also been used in evaluation of the potential that components of the infection process can produce complete extinction of pathogens populations before a systemic infection (e.g., establishment of a steady pathogen population *in vivo*) occurs (e.g., Blaser and Kirschner, 1999; Coleman and Marks, 2000). These studies are described in the following section.

C.1.3. Mechanistic and Physiologically-Based Models of Infection

Models of the infection process (i.e., mechanistic dose-response models) may be developed with varying degrees of resolution. These models differ in the components of the infection process that are explicitly modeled and whether they are deterministic or stochastic. Early attempts at developing mechanistic dose-response models focused on stochastic pathogen birth and death processes or on division of the infection process into stages that might be modeled separately. Under the assumption that pathogens divide and are removed (via innate or active immune system processes or other means) at constant rates, μ and λ, Bailey (1964) developed expressions for the probability that an *in vivo* population of size N is realized at time t:

$$p_N(t) = \sum_{j=0}^{\min(d,N)} \binom{d}{j} \binom{d+N-j+1}{d-1} A^{d-j} B^{N-j} (1-A-B)^j \qquad \text{[C-6]}$$

and for the probability that the pathogen population reaches extinction at time t:

$$p_0(t) = A^d \qquad \text{[C-7]}$$

or as $t \rightarrow \infty$:

$$\lim_{t \to \infty} p_0(t) = \left(\frac{\mu}{\lambda}\right)^d \quad \text{[C-8]}$$

In Equation C-6,

$$A = \frac{\mu\left[1 - e^{(\mu-\lambda)t}\right]}{\lambda - \mu e^{(\mu-\lambda)t}} \quad \text{[C-9]}$$

$$B = \frac{\lambda\left[1 - e^{(\mu-\lambda)t}\right]}{\lambda - \mu e^{(\mu-\lambda)t}} \quad \text{[C-10]}$$

Morgan and Watts (1980) used Equation C-6 to derive an expression for the probability that a single pathogen (d = 1) achieves a threshold population at time t:

$$p(n \geq N; t) = (1 - A)B^{N-1} \quad \text{[C-11]}$$

Alternately, the probability that the incubation period, T, is less than a time, t, is:

$$p(T \leq t; t) = \frac{(1 - A)B^{N-1}}{1 - \mu/\lambda} \quad \text{[C-12]}$$

Morgan used MLE to fit the dose-response model given in Equations C-9 to C-11 to incubation period data drawn from a study of the incubation period of streptococcal sore throat (Sartwell, 1950). Morgan's estimates for growth rate, λ, death rate, μ, and number of organisms present *in vivo* at the incubation time, N, were 0.236/hour, 0.190/hour$^-$ and 46.51, respectively. Morgan hypothesized that the very low value estimated for N results from neglecting complications such as site heterogeneity, eclipse periods and hosts differing in response due to natural or acquired resistance, age, and allergic states.

Williams (1965) derived an expression for the probability of a dose of d organisms achieving a net population of N or more organisms at time t. Note the difference between one organism giving rise to a population of N organisms in time t and d organisms achieving a net population of N organisms at time t. Assuming a constant birth rate, λ, and death rate, μ, Williams showed that the distribution of incubation periods (time to achieve a net population of N organisms) for an inoculum of dose, d, is given by:

$$f(\tau) = \frac{\sqrt{d}}{e^d - 1} \exp\left(-\tfrac{1}{2}\tau - e^{-\tau}\right) I_1\left(2d^{1/2}e^{-\frac{1}{2}\tau}\right) \quad \text{[C-13]}$$

In Equation C-13, $\tau = (\lambda - \mu)\, t - \ln[N(1 - \nu)]$, $\nu = 1 - \mu/\lambda$, and I_1 denotes a first Bessel function of the imaginary argument. The model presented as Equation C-13 was found to provide an excellent fit to the distribution of incubation periods observed in an outbreak of streptococcal sore throat associated with consumption of contaminated milk, although details of the calculations were not provided.

Brookmeyer et al. (2005) developed a time-dependent dose-response model referred to elsewhere as a competing risks model (Gutting et al., 2008). One of the authors' stated motivations in developing a mechanistic model for *Bacillus anthracis* infection was utilization of available data on infection by *B. anthracis* (spore germination rates, clearance rates, growth rates) in the absence of detailed human dose-response data developed in experimental studies. Assuming a constant risk of spore germination per unit time, ω, and a constant risk per unit time of clearance from the lung, κ, and assuming that spore germination implies systemic infection, Brookmeyer et al. (2005) showed that the cumulative attack probability for inhalation anthrax may be estimated as:

$$F(t) = 1 - \exp\left[\frac{-d\omega}{\omega+\kappa}\left(1 - e^{-(\omega+\kappa)t}\right)\right] \qquad \text{[C-14]}$$

Inspection of Equation C-14 shows that the Brookmeyer competing-risks model yields the exponential dose-response model in the limit $t \rightarrow \infty$.

Blaser and Kirschner (1999) developed a deterministic model for *in vivo* pathogen growth, including immune system response. In that study, stocks and flows of five quantities—mucus-living *Helicobacter pylori*, *H. pylori* attached to epithelial cells, density of bacterial nutrients released via inflammation, density of effector molecules, and host response—were included in a system of ordinary differential equations describing the dynamics of these quantities. In mucus, the conservation equation for *H. pylori* accounted for growth (first-order with respect to nutrient availability), loss due to mucus shedding and migration, and gain due to emigration. The conservation equation for *H. pylori* on epithelial cells included a growth term, a loss term related to sloughing, and terms accounting for immigration and emigration. The authors used their model to explore the importance of the parameters in their model of infection, determining that the parameter that describes the ability of the immune system to respond was the most important determinant of whether there would be extinction (all pathogens are removed from the system) or whether sustained growth occurs and that the bacteria growth parameter had a limited effect on the ability of pathogens to initiate infection but was the most important factor in determining the time required for pathogens to reach a steady population *in vivo*. In subsequent modeling work, Blaser and Kirschner (2007) used a deterministic model to explore infections with slow progression or latent periods, during which there is equilibrium between host response and pathogen population dynamics. Taken together, the two studies by Blaser and Kirschner (1999, 2007) demonstrate the utility of deterministic models in exploring complex infection processes.

Coleman and Marks (2000) developed both stochastic and deterministic models of non-typhoid salmonellosis and used the models to identify factors that influence the shape of the dose-response curve in the low-dose region. In that study the important events occurring in the course of *Salmonella* infection were posited to be survival of ingested bacteria to the target, colonization, engulfment, intracellular survival, migration and multiplication, damage, and AGI. The authors suggested stochastic models for each of these processes and presented an alternative formulation based on a predator-prey framework. As pointed out by Coleman and Marks (2000) for infection by non-typhoid *Salmonella* and also by Levin and Antia (2001) for infections in general, there may be physiological and biological process that do not conform to the assumptions underlying the beta-Poisson or exponential dose-response model, including clumping of pathogens in the ingested

dose, quorum sensing, and the possibility that organisms do not exhibit independent action. In the context of MRA model results and results of feeding studies of healthy adult human volunteers, the authors made a case for sub-linearity of response at low dose. The authors note that the potential for sub-linear low-dose-response is likely differ between pathogen-host combinations and that additional data such as *in vitro* studies may provide information for parameter selection for mechanistic infection models. Development and validation of additional mechanistic models for infection provides an avenue for evaluating low-dose-response.

The inherent variability of host-pathogen processes suggests use of stochastic models for describing *in vivo* processes leading to infection. Allen and Allen (2003) describe Markov-chain and stochastic differential equation models for estimating the pathogen burden *in vivo* as a function of time. Their model of the infection process is relatively simplistic, comprised only of birth and death processes in which birth and death rates may vary with time or pathogen density, but their framework is amenable to inclusion of additional components (e.g., immune system components, pathogens in different states). Based on evaluation of different models for a relatively simple case, the authors concluded that combinations of deterministic and stochastic models offer the greatest opportunity for including relevant features of the infection process in a computationally tractable framework.

Recent dose-response modeling efforts have included development of highly-detailed, physiologically-based models of the infection process. In their recent assessment of anthrax dose-response models, Gutting et al. (2008) outlined the components of a hypothetical physiologically-based biokinetic model of infection and response to aerosols of *Bacillus anthracis.* In the model, the fate and transport of *B. anthracis* spores and vegetative cells is tracked in regions of the respiratory system, in macrophages, in the blood and in lymph nodes. As done by Brookmeyer et al. (2005) in their development of a competing risks model for inhalation anthrax, Gutting et al. (2008) estimate model parameters for use in their biokinetic model using physiological and microbiological data not collected in quantal dose-response studies or epidemiological investigation. However, details of the techniques used for parameter estimation or of the model were not provided in the study by Gutting et al. (2008).

C.2. Use of Bayesian Methods in Microbial Risk Assessment

Bayesian methods are being increasingly used by several researchers in microbial risk assessment to estimate dose-response model parameters. In general, a dose-response function gives the probability of illness or infection as a function of the dose and of several unknown parameters. Experimental data are collected from subjects accidentally (in an outbreak) or deliberately (in a controlled experiment with volunteer human subjects or with animal subjects) exposed to a microbial dose that can be measured or estimated. The numbers of subjects that become infected or ill for each dose level are observed, leading to a binomial likelihood. That is, the probability of n "successes" out of N trials of dose level d, where "success" means illness or infection and the success probability is given by the dose-response function. The "traditional" frequentist statistical approach uses the binomial likelihood only, and chooses parameter values to maximize the likelihood.

In general, if N_i subjects are exposed to a mean dose D_i, and n_i of them developed infection, then the likelihood for the full population (all dose groups) is given by the following:

$$\prod_{i=1}^{\#\text{doses}} \frac{N_i!}{N_i!(N_i - n_i)!} \times [P(\text{infection} \mid D_i)]^{n_i} \times [1 - P(\text{infection} \mid D_i)]^{N_i - n_i} \qquad \text{[C-15]}$$

The dose-response function is the function $P(\text{Infected} \mid D)$, which will depend upon the mean dose D as well as the unknown parameters.

To estimate the uncertainty of the estimated dose-response parameters and dose-response function, 95% confidence intervals can be calculated using standard asymptotic theory, valid when the sample sizes ("n") are large. The asymptotic theory uses the likelihood function (Equation C-15) to derive an estimated standard error for each parameter, and the 95% confidence interval can then be estimated as the maximum likelihood estimate plus or minus 1.96 standard errors. The 1.96th is the 97.5th percentile of a standard normal distribution, which applies because for large samples the estimated parameter approximately has a normal distribution.

Alternatively, and preferably for the small sample sizes usually available in microbial risk assessment, a Monte Carlo bootstrap resampling method can be used to estimate the uncertainty by randomly sampling with replacement from the original data and fitting the model to each of the resampled data sets. Bayesian methods are preferable to bootstrapping with smaller sample sizes.

Bayesian methods exploit available subjective and related information in addition to the numeric data from the experiment or outbreak. Ideally, the investigator expresses their initial assessment of the unknown parameter distribution, prior to examining the data, by defining a prior probability distribution for the parameters. The prior probability distribution is defined based on subjective information and professional judgment.[27] Using Bayes' rule, the posterior probability distribution for the parameters given the data can be calculated. From Bayes' rule, the posterior distribution equals the prior distribution for the parameters multiplied by the likelihood for the data (given the parameters) and then divided by a normalizing constant. The normalizing constant is the integral of the product of the prior and likelihood over all possible parameter values.[28] In a Bayesian analysis, uncertainty intervals for the parameters and the dose-response function can be calculated from the posterior distribution as "credible intervals"; a 95% credible interval has a 95% probability of including the parameter value, given the data.

[27] Some Bayesian researchers use a more objective approach called the empirical Bayes method that is based on an hierarchical model such that the likelihood depends upon parameters that have distributions depending upon other parameters, called hyperparameters. A frequentist approach, such as maximum likelihood, is used to estimate the hyperparameters and thus estimates the prior distribution without the use of subjective information. At this time, the authors are not aware of any applications of empirical Bayes methods to microbial risk assessment.

[28] Suppose θ is the vector of unknown parameters, and has a prior distribution with probability density function $f(\theta)$. Suppose the data X has a likelihood given by $g(X \mid \theta)$, for example, Equation C-15. Then the posterior distribution will have a probability density given by $f(\theta)\, g(X \mid \theta) / k(X)$, where $k(X)$ is the normalizing constant. This is Bayes' rule. The normalizing constant is the integral

$\int f(\theta)\, g(X \mid \theta)\, d\theta$, integrated over all possible values of θ.

The constant $k(X)$ does not depend upon the parameters although it will depend upon the data X.

The choice of a suitable prior distribution is crucial and can be controversial. Recent published MRAs have usually had very little subjective information to rely on for choosing a prior distribution and the investigators have chosen a "non-informative" prior distribution to represent the lack of prior information. The researchers have usually published their choice of non-informative prior, but have not usually provided a rationale for their choice over other possible non-informative priors.[29] For example, Teunis and Havelaar (2000) used a beta-Poisson model, described below, and chose the prior distribution for their parameters α and β such that their logarithms (base 10) were assumed to have a wide uniform distribution from -12 to +6 and the parameters were assumed independent. Englehardt (2004), using the same beta-Poisson model, instead chose a joint uniform prior distribution for the parameters α and β. Teunis et al. (2004) also used a beta-Poisson model, but used another non-informative prior, such that $\alpha/(\alpha+\beta)$ is uniform from 0 to 1 and $\log_{10}(\alpha+\beta)$ is normally distributed with mean 0 and standard deviation 10. If the non-informative prior distribution is wide then the posterior probability distribution should not be sensitive to the choice of non-informative prior, which justifies the name "non-informative prior." However, researchers have used different non-informative priors for the same model, which suggests that the choice of the so-called non-informative prior can influence the results.

In the past, Bayesian researchers were much more limited in their choice of prior distributions because they needed to choose a distribution to make the calculations tractable (a "conjugate" prior), particularly the calculation of the normalizing constant. More recently, with MCMC methods and fast computing methods, the calculations can be easily executed for a much wider variety of prior distributions using Monte Carlo simulation methods.

The MCMC method describes a group of methods used to simulate values from a probability distribution for which direct analytical calculations are difficult, intractable, or inconvenient. Gilks et al. (1996) provide a good description of these methods. Well-validated software packages are available to perform these calculations, including WinBUGS and Mathematica. For Bayesian MCMC analyses, the simulated probability distribution is the joint posterior distribution of the parameters given the data. Thus, at each step of the Markov chain, a vector of parameter values is simulated, rather than a single parameter value. Furthermore, it is unnecessary to know the normalizing constant for the posterior distribution, which is often the most difficult part of the calculation. All that is needed are some constant multiples of the prior distribution and the likelihood. The normalizing constants needed to make the prior and likelihood integrate to one are not needed. A version of the Metropolis-Hastings algorithm (Hastings, 1970; Gilks et al., 1996) is used at each step to simulate from the posterior distribution without knowing the normalizing constant.[30] Instead of being statistically independent, the consecutive values form a Markov chain,

[29] Teunis et al. (2004, 2005, 2008a,b) transformed their parameters using logarithms and logit functions to avoid high corrections between the parameters and thus improve the estimation. However, this does not really explain their choice of non-informative prior for the transformed variables.

[30] Suppose that the product of the prior and likelihood is equal to K x $f(\theta)$, where θ is the vector of all the unknown parameters and K is an unknown normalizing constant that will depend upon the data values; that is, for MRA, the numbers of illnesses or infections observed. To obtain the posterior distribution, K can, in principle, be calculated as the reciprocal of the integral of f over the range of possible parameter values, but this calculation is often very difficult analytically. Let $q(\theta, \varphi)$ be any chosen proposal distribution, which is a probability density for the next parameter vector φ that may depend upon the previous parameter vector θ. If the parameter vector at the previous step is θ, one must first randomly sample a parameter vector from $q(\theta, \varphi)$ to obtain a candidate vector φ^*. With probability α, one can accept the candidate vector, so that the vector at the next step of the Markov chain is φ^*. With

so that the statistical distribution for one value depends upon the previous value. Using the MCMC method, the Markov chain has a limiting, stationary distribution, so that after a sufficiently long "burn-in" period the values have the desired probability distribution.[31]

The articles reviewed for this discussion do not specify the details of the Metropolis-Hastings algorithms used. Many Bayesian analysts use the Gibbs sampler, which is a special version of the Metropolis-Hastings algorithm that always has acceptance probability 1, so that a new parameter vector is selected at each step. Instead of jointly updating all the parameters in a single step, the Gibbs sampler simulates each of the parameters in turn.[32] An algorithm such as Adaptive Rejection Sampling (Gilks and Wild, 1992) is used to generate samples from the distribution of each parameter without needing to calculate the normalizing constant.[33]

Bayesian modeling has been used by MRA researchers in various ways. Several authors have used both Bayesian and frequentist (likelihood-based) methods (Teunis and Havelaar, 2000; Messner et al., 2001). Often the frequentist approach is used to provide maximum likelihood estimates of the dose-response function and the Bayesian approach is used to calculate uncertainty intervals (e.g., 80 or 95% credible intervals for the parameters or the dose-response). The frequentist likelihood ratio test is used to compare different dose-response models. Several approaches use the mode of the Bayesian posterior distribution to select the dose-response function (Teunis et al., 2004, 2005, 2008a,b). The posterior mode is given by the parameters that maximize the posterior

---

probability 1- α, one rejects the candidate vector, so that the vector at the next step of the Markov chain is, again, θ. The probability α is calculated as $\min(f(\varphi^*)\, q(\varphi^*, \theta)/\{f(\theta)\, q(\theta, \varphi^*)\}, 1)$. Because f appears in both the numerator and denominator, the unknown K cancels out and is not needed. A good choice of the proposal distribution will have high acceptance rates and fast convergence to the stationary distribution.

[31] A burn-in period of about 5000 steps is usually sufficiently long that the stationary distribution has been reached; various convergence tests can be used to assess convergence. Values generated during the burn-in period are discarded, and one usually selects every k^{th} value after the burn-in period for some suitably large k (e.g., 10, 20, 100) so that the remaining "thinned" sequence of values are approximately independent. Thus, the thinned values after the burn-in period can be treated as if they were a random sample from the given probability distribution.

[32] Suppose that there are n unknown parameters in the posterior distribution. Instead of generating a new multivariate vector of n parameters from a joint distribution, the Gibbs sampler generates each parameter in turn from the univariate "full conditional" distribution of that parameter given the values of all the other parameters and the data. Thus, each Markov chain step becomes a sequence of sub-steps where the n parameters are scanned in turn and the m^{th} parameter value is randomly selected from the conditional distribution of the m^{th} parameter given the data and the most current values of the remaining n-1 parameters (i.e., the values of the first m-1 parameters from the current scanning steps and the values of the last n-m parameters from the previous Markov chain vector). An algorithm such as Adaptive Rejection Sampling (Gilks and Wild, 1996) is used to generate samples from each full conditional distribution without needing to calculate the normalizing constant.

[33] Suppose that the product of the prior and likelihood is equal to $K(\theta_{-m}) \times g(\theta_m, \theta_{-m})$, where θ_m is the m^{th} unknown parameter, θ_{-m} is the vector of the other n-1 unknown parameters, and $K(\theta_{-m})$ is an unknown normalizing constant that will depend upon the data values and the values of the remaining n-1 parameters. To obtain the full conditional distribution, $K(\theta_{-m})$ can, in principle, be calculated as the reciprocal of the integral of g over the range of possible values of θ_m, treating the other parameters as constants, but this calculation is often very difficult analytically. Adaptive rejection sampling (Gilks and Wild, 1992) randomly generates values of θ_m from g, without knowing the normalizing constant. The method requires that the function g is log-concave in θ_m, which holds for many distributions if the parameters are appropriately defined. The method may need to generate and reject several random values until the final value is accepted, but at each rejection, more exact bounds for g are calculated so that the probability of future rejections rapidly decreases. If the full conditional distribution is not log-concave, then the Metropolis-Hastings algorithm can instead be used to generate values from the full conditional distribution without knowing the normalizing constant.

probability, defined as the product of the prior and the likelihood; it is again not necessary to calculate the normalizing constant.

Englehardt and Swartout have published several papers (Englehardt, 2004; Englehardt and Swartout, 2004, 2006, 2008) advocating the use of the predictive Bayesian approach, which is the unconditional dose-response probability, calculated as the integral of the posterior distribution multiplied by the dose-response function, integrated over the parameter space. This can be thought of as the dose-response function averaged over the uncertainty distribution. The predictive Bayesian method has the advantage of producing an estimated dose-response function that is more protective of public health than the maximum likelihood estimate, because at low doses the estimated risk is generally higher. The method also has the advantage of avoiding the need to specify a frequentist confidence level or a standard Bayesian prediction interval probability level, which avoids potential inconsistencies when comparing risks from different health stressors; the more risky stressor can depend upon the probability level chosen. On the other hand, upper bounds of confidence or prediction intervals can be thought of as estimating the risk under a "worse-case" scenario, and regulators may prefer a worse-case scenario approach to the predictive Bayesian approach that represents the average scenario (averaging estimates over the parameter uncertainty).

C.2.1. Comparison of Bayesian and Frequentist Methods

Before Bayesian methods were applied to MRA, risk assessors were generally limited to simpler model formulations and approximate uncertainty estimates. Risk assessors also could not take advantage of any available subjective information on the values of the unknown parameters. An advantage of the Bayesian approach over the frequentist approach is the ability to incorporate prior information, although for the MRAs in the current literature this is not very important because the prior information is too limited and so non-informative priors have been used. A more important advantage is that the uncertainty intervals from a Bayesian analysis are easier to interpret and are usually not interpreted incorrectly—a Bayesian 95% credible interval for the dose-response is interpreted as having a 95% probability of including the true probability of illness (or infection) given the available data. The risk (probability of illness or infection) is treated as being random. A frequentist 95% confidence interval is properly interpreted as having a 95% probability of including the true probability of illness (or infection) in an identical future experiment, so that 95% of a large number of identical future experiments will give confidence intervals that include the true risk. The risk is treated as being an unknown constant. Lay persons will very often incorrectly interpret the confidence interval as if it had the same meaning as the Bayesian credible interval.

Bayesian dose-response uncertainty calculations using MCMC also have the advantages of being easier and more exact than the frequentist confidence intervals. Because the dose-response function is a complicated function of multiple parameters, the confidence intervals are hard to calculate or approximate analytically. The bootstrap or similar Monte Carlo resampling methods can avoid these difficult analytical calculations but this often requires more computation than MCMC. Furthermore, the large sample theory estimates of the confidence intervals are poor approximations for the small samples typically found in MRA. While bootstrap estimates are more reliable for small samples, they are also approximations to the true uncertainty distributions, even if the number of bootstrap simulations is tending to the infinite limit. The MCMC uncertainty

estimates are exactly correct for the posterior distribution assuming that the burn-in period is sufficiently long that the chain can be considered stationary (ignoring the imperfect nature of computer random number generation).

A further major advantage of the Bayesian method for MRA is the ability to use a hierarchical Bayesian model to model cases where the host or pathogen response parameters vary over the population of humans or organisms (e.g., see Messner et al., 2001, discussed below). This type of meta-analysis is easier to apply in a Bayesian framework.

The major disadvantage of the Bayesian approach is the requirement for developing a prior distribution that, in principle, is subjective and thus depends on the information available to the investigator. Different investigators can choose different priors for the same model formulation, even if the prior are "non-informative." The subjective nature of a prior distribution can be disturbing. On the other hand, Bayesian statisticians often point out that the investigator's choice of dose-response function or other mathematical model is also a subjective choice.

C.2.2. Applications of Bayesian Methods to Microbial Risk Assessment

In one of the earliest Bayesian analyses of microbial risks, Teunis and Havelaar (2000) modeled rotavirus, *Campylobacter*, and *Vibrio cholerae* dose-response data using a series of models. In the Poisson model, an individual is exposed to a number of organisms (e.g., colony forming units [CFU]) that is assumed to have a Poisson distribution with a mean dose D equal to the volume ingested multiplied by the average number of CFU per unit volume. Each single organism independently has the same hit probability (r) of infecting the subject. It follows that the probability of infection at dose D is exponential with parameter $r \leq 1$:

$$\text{Prob (infected} \mid D, r) = 1 - e^{-rD}. \qquad \text{[C-16]}$$

In the beta-Poisson model, r is assumed to vary among hosts or organisms with a beta distribution with parameters $\alpha > 0$ and $\beta > 0$. This gives the dose-response function as follows:

$$\text{Prob (infected} \mid D, \alpha, \beta) = 1 - {}_1F_1(\alpha, \alpha + \beta, -D). \qquad \text{[C-17]}$$

The function ${}_1F_1$ is the Kummer confluent hypergeometric function.

The gamma-Poisson model is an approximation to the beta-Poisson model of the following form:

$$\text{Prob (infected} \mid D, \alpha, \beta) = 1 - (1 + D / \beta)^{-\alpha} . \qquad \text{[C-18]}$$

This model is also obtained if either r has a gamma distribution that includes all values $r > 0$ (even though the hit probability r cannot exceed 1), or, more realistically, if the density has a gamma distribution. The model in Equation C-18 was originally called the beta-Poisson model, because it was derived as an approximation to the exact beta-Poisson model in Equation C-17. The same three equations can be used to model the probability of illness, particularly for outbreaks or studies, where the numbers of infection cases are not reported.

Teunis and Havelaar (2000) fitted all three models and estimated the parameters r, α, and β using maximum likelihood (i.e., choosing values to maximize the probability of the data given the parameters, which is a product of binomial probabilities). The likelihood is given by Equation C-1 above. Using the maximum likelihood estimates of the parameters in the dose-response function (Equations C-1, C-2, or C-3) provides the maximum likelihood estimate of the dose-response function. They also estimated approximate 95% confidence intervals and regions for the parameters using the likelihood.

The authors also used a bootstrap resampling method to estimate the uncertainty of the dose-response function by randomly sampling with replacement from the original data and fitting the dose-response model to each of the resampled data sets. Thus, each bootstrap sample gives a different dose-response function.

Teunis and Havelaar (2000) used a Bayesian MCMC approach primarily to more easily compute the uncertainty estimates and to compare their maximum likelihood estimates with Bayesian estimates. Their prior distribution for the parameters α and β assumes they are independent and that their logarithms (base 10) have a uniform distribution from -12 to +6. The MCMC method was used to generate pairs of parameter values α and β from the posterior distribution. They found that the likelihood-based confidence regions for the parameters probability matched well to the sampled Bayesian posterior distribution. For each parameter pair, the dose-response function was calculated. For each dose D, a 95% credible interval for the probability of being infected is given by the 2.5^{th} to 97.5^{th} percentiles of the set of dose-response functions evaluated at dose D.

Messner et al. (2001) used Bayesian methods to analyze the results of three human volunteer studies, each using different isolates of *Cryptosporidium*—IOWA, TAMU, and UCP. For each individual study, they fitted an exponential model (Equation C-2) by maximum likelihood and then computed the maximum likelihood estimate of the dose-response function. They compared their results with a Bayesian analysis based on the assumption that ln(r) has a uniform distribution over the entire real line.[34] The means and medians of the Bayes predictive distributions were very similar to the maximum likelihood estimates. Given the assumed distribution for log(r), its posterior density is proportional to the likelihood function.

To combine results from the three studies in a meta-analysis, Messner et al. (2001) used a hierarchical Bayes model that had several groups of parameters. At the first level, the hyperparameters are parameters with a prior distribution that does not depend on any other parameter. At the second level, some parameters are assigned distributions that depend upon the

[34] Such a prior is not a proper probability distribution because it cannot integrate to 1, but in many cases an improper prior can be used to calculate a valid posterior distribution. This improper uniform prior can be regarded as being the limit of a uniform distribution for log(r) over the range -M,M as M tends to infinity.

values of the hyperparameters. At the third level there are parameters that have distributions that depend upon the first and/or second level parameters. The hierarchy can have multiple levels, although most applications to MRA have at most two levels.

Messner et al. (2001) defined two hyperparameters μ and σ. Their prior distributions were not listed in their paper. The parameters r for each study were assumed to be independently drawn from a normal distribution with mean μ and standard deviation σ. This normal distribution represents variability of r between isolates, i.e., the probability of infection from a dose of a single organism depends upon the isolate. Thus, the model has a total of five parameters. The MCMC method using the Gibbs sampler was used to generate samples of parameter vectors from the posterior distribution given the data from all three studies. Eighty percent credible intervals (10^{th} to 90^{th} percentile of the posterior distribution) were thus calculated for the parameters and for the dose-response function at a dose of one oocyst.

Teunis et al. (2004) used Bayesian modeling to analyze data from an outbreak of *E. coli* O157:H7. They modeled the dose-response functions using the beta-Poisson model shown in Equation C-17. A non-informative prior was selected such that $u = \alpha/(\alpha+\beta)$ is uniform from 0 to 1 and $v = \log_{10}(\alpha+\beta)$ is normally distributed with mean zero and standard deviation 10. The transformed parameter u is the mean of the beta distribution for r, and v is inversely related to the variance of the beta distribution. The parameters u and v are assumed independent. The transformation improves the parameter estimation since if there is only a single dose value, α and β are highly correlated. The parameter values corresponding to the mode of the posterior distribution were calculated directly by numerically maximizing the posterior probability, which is the same as maximizing the product of the prior distribution and the likelihood. The posterior probability equals the prior multiplied by the likelihood and divided by the normalizing constant, which does not depend upon α and β. Using the posterior mode parameter values in Equation C-17 gives the posterior mode dose-response equation. The uncertainty of this dose-response function was characterized using MCMC sampling of parameter vectors. The dose-response function was calculated for each parameter vector and the percentiles of the response probability for each dose were plotted. Frequentist likelihood ratio tests were used to compare different dose-response models.

Teunis et al. (2005) analyzed *Campylobacter jejuni* dose-response using Bayesian methods. Data from both a human volunteer study and two outbreaks caused by drinking raw milk (children and teachers visiting farms; one in Holland and one in the UK) were combined in this analysis. The model incorporated both the probability of infection and the conditional probability of illness given infection for children. First, for the outbreak a certain probability of illness (p0) was assumed for those who were unexposed to the raw milk but might have become ill due to an alternative route of transmission. Second, a beta-Poisson model was used to model the probability of infection given a mean dose (D). Third, a model for the conditional probability of illness, given that the individual is infected and had mean dose D, was developed as follows:

$$\text{Prob}(\text{ill} \mid \text{infected}, D, r, \eta) = 1 - (1 + \eta D)^{-r}. \qquad \text{[C-19]}$$

Non-informative prior distributions for the parameters were defined by assuming that all parameters are independent and that $\text{logit}(\alpha/(\alpha+\beta))$, $\log_{10}(\alpha+\beta)$, $\log_{10}(r\eta)$, $\log_{10}(r/\eta)$, and $\text{logit}(p0)$ are all normally distributed with mean 0 and standard deviation 10. By definition $\text{logit}(x) =$

log(x/(1-x)). The posterior mode parameter values were calculated by directly maximizing the posterior probability. These values were used to compute the posterior mode dose-response functions for the probability of infection (Equation C-17) and the probability of illness given infection (Equation C-19). Uncertainty intervals for these dose-response functions were computed by using MCMC to simulate vectors of parameter values.

Teunis et al. (2008a) analyzed data from eight outbreaks of *E. coli* O157:H7 using a hierarchical Bayes model. A homogeneous exposure model used a beta-Poisson dose-response function for the probability of illness (Equation C-17). A heterogeneous version of the exposure model also included known values of a dispersion parameter, treated as being the shape parameter for a gamma distribution of the microbial densities. Using a different notation to that used in the paper, the hyperparameters m1, m2, s1, and s2 were assumed to be independent and have distributions such that m1 and m2 were normally distributed with mean -8 and standard deviation 8, and s1 and s2 were gamma (0.001,1000) distributed. For outbreak i, $\text{logit}(\alpha(i)/(\alpha(i)+\beta(i)))$ and $\log_{10}(\alpha(i)+\beta(i))$ were assumed independently normally distributed with means m1 and m2 and standard deviations s1 and s2. To obtain an overall group dose-response function, representing the dose-response function for a future random outbreak, the α and β parameters of that outbreak were assumed to be generated from the prior distribution of the hyperparameters; that is, $\text{logit}(\alpha/(\alpha+\beta))$ and $\log_{10}(\alpha+\beta)$ were assumed to be independently normally distributed with means m1 and m2 and standard deviations s1 and s2.

Teunis et al. (2008a) fitted these models using MCMC. The posterior mode dose-response function for each outbreak was estimated by finding the sample parameter vector with the highest value of the joint posterior (partial) probability, which is the product of the prior density for the hyperparameters, the conditional density for α(i) and β(i) given the hyperparameters, and the likelihood for the outbreak i. The overall estimates of the dose-response function for a future outbreak were estimated by sampling α and β from the prior distribution of the hyperparameters and computing the dose-response function for each pair. This gives a set of dose-response functions. For each dose, the percentiles of the probability of being ill were computed and plotted as a contour plot.

Teunis et al. (2008b) used Bayesian methods to analyze dose-response functions for the Norwalk virus based on a volunteer study. Similar methods to the above studies were employed so the details are not discussed here.

Englehardt (2004) compared maximum likelihood methods to a predictive Bayesian dose-response approach. The method was applied to rotavirus data. First, he discussed the likelihood-based Benchmark Dose Method, which computes a confidence interval for the dose at which a certain change or percentage change in risk occurs and defines the benchmark dose as the lower confidence value. He pointed out that for two different health stressors, it is possible that the dose-response curves can intersect, in which case the more risky stressor (the stressor with the least benchmark dose) depends upon the confidence level chosen. A similar issue arises with Bayesian analyses using credible intervals to account for uncertainty; that is, the results depend upon the assumed "confidence" level. Englehardt recommends averaging the dose-response function over the posterior distribution of the parameters. In other words, the predictive Bayesian dose-response model for dose D is calculated as the integral of the posterior distribution of the parameters given

the data multiplied by the dose-response function at dose D. This integral is over the entire probability space. The predictive Bayesian dose-response is the unconditional dose-response function and can be thought of as the dose-response function averaged over the uncertainty distribution. This can be compared to frequentist or more standard Bayesian approaches for which upper bounds of confidence or prediction intervals can be thought of as estimating the risk under a "worse-case" scenario. Regulators may prefer a worse-case scenario approach to the predictive Bayes approach, which represents the average scenario.

Englehardt (2004) applied the predictive Bayes approach to rotavirus data using the beta-Poisson model. The maximum risk for any dose D is calculated using the exponential model with r = 1, which assumes a hit probability equal to 1 so that infection is guaranteed (100%) if any organisms are ingested. The minimum risk for any dose D is assumed to be obtained from the maximum likelihood estimates of the parameters α and β. The predictive Bayes dose-response function has a risk between the minimum and maximum risk. Englehardt (2004) points out that in general, if enough data are available, then at low doses, the predictive Bayes risk will be lower than the observed risk (proportion of illnesses), so that the approach is conservative (health-protective) compared to maximum likelihood methods. To calculate the posterior distribution, Englehardt used the MCMC method assuming an improper uniform prior for α and β. It is not clear from the paper how the integral of the posterior multiplied by the dose-response function was calculated. Although direct numerical integration is possible, in principle, a reasonable approach would be to compute the dose-response function (probability of illness) for each dose D and each sampled pair of parameter values, and then average the probability of illness over the entire sample. Note that Englehardt defines the normalizing constant k as the constant that normalizes the likelihood function. This is not correct in general since the normalizing constant k normalizes the posterior distribution (i.e., the product of the prior and the likelihood).

Englehardt and Swartout (2004) applied the predictive Bayes approach to the *Cryptosporidium parvum* data analyzed by Messner et al. (2001). However, these analyses separated out the results for subjects with Ab+ and Ab- serum-antibody status. First, maximum likelihood estimates for the beta-Poisson models were computed for each study (isolate) and Ab+ or Ab- status. Second, a representative population of sensitive, Ab+, and Ab- subjects was simulated using the maximum likelihood fitted dose-response functions; sensitive subjects were assumed to always respond at the doses tested. Each simulated population is a parametric bootstrap sample. A beta-Poisson model was fitted to each bootstrap sample using maximum likelihood. The set of maximum likelihood estimates was used to compute 95% confidence intervals for the probability of infection for each strain.

Englehardt and Swartout (2004) computed a predictive Bayes distribution for the r for a random isolate and for the dose-response function. For a random isolate, r is assumed to have a beta distribution with parameters α and β, assigned a joint uniform prior. The likelihood of the three observed r values is given by the product of three beta distributions; each observed r value is the mean for the bootstrap simulations of that isolate. Thus, the predictive Bayes distribution for r is defined by multiplying the beta distribution for r by the posterior probability for α and β given the three sampled values of r, and then integrating over the parameter space. This is the marginal distribution of r. To obtain the predictive Bayes dose-response function, the marginal probability

for r was multiplied by the exponential dose-response function (Equation C-2) and integrated over $0 \leq r \leq 1$.