Quantitative Microbial Risk Assessment:

Application for Water Safety Management

World Health Organization

WHO Library Cataloguing-in-Publication Data

Quantitative microbial risk assessment: application for water safety management.

1.Water Microbiology. 2.Water Quality. 3.Waterborne Diseases – prevention and control. 4.Risk Assessment. I.World Health Organization.

ISBN 978 92 4 156537 0 (NLM classification: WA 675)

© World Health Organization 2016

All rights reserved. Publications of the World Health Organization are available on the WHO website (http://www.who.int) or can be purchased from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; email: bookorders@who.int).

Requests for permission to reproduce or translate WHO publications – whether for sale or for non-commercial distribution – should be addressed to WHO Press through the WHO website (http://www.who.int/about/licensing/copyright_form/index.html).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

Design and layout: Paprika (Annecy, France)

Printed by the WHO Document Production Services, Geneva, Switzerland

Quantitative Microbial Risk Assessment:

Application for Water Safety Management

World Health Organization

TABLE OF CONTENTS

ACKNOWLEDGEMENTS .... **IX**

ABBREVIATIONS .... **XI**

EXECUTIVE SUMMARY .... **XIII**

1 INTRODUCTION: THE EVIDENCE BASE FOR ADEQUATE RISK MANAGEMENT **1**

1.1 Moving towards preventive, risk-based water safety management 1

1.2 Purpose of this document 2

1.3 Intended audience and reading guide 3

1.4 Structure of this document 3

2 SPECTRUM OF RISK ASSESSMENT APPROACHES **5**

2.1 Sanitary inspection 7

2.1.1 Strengths *10*

2.1.2 Limitations *10*

2.2 Risk matrix 10

2.2.1 Strengths *12*

2.2.2 Limitations *12*

2.3 QMRA 12

2.3.1 Problem formulation *13*

2.3.2 Exposure assessment *13*

2.3.3 Health effects assessment *14*

2.3.4 Risk characterization *14*

2.3.5 Strengths *16*

2.3.6 Limitations *17*

2.4 Selecting the appropriate risk assessment approach 17

3 FRAMEWORK FOR WATER-RELATED QMRA **21**

3.1 Problem formulation 22

3.2 Exposure assessment 22

3.3 Health effects assessment 22

3.4 Risk characterization 23

4 VARIABILITY AND UNCERTAINTY 25
4.1 Variability 25
4.2 Uncertainty 26
4.3 Accounting for variability and uncertainty in QMRA 29

5 PROBLEM FORMULATION 33
5.1 Interaction between risk manager and risk assessor 33
5.2 Fit-for-purpose 34
5.3 Including risk management options in the QMRA 35
5.4 Defining the scope of the QMRA 35
5.4.1 Hazard identification 35
5.4.2 Exposure pathways 36
5.4.3 Health outcomes 36

6 EXPOSURE ASSESSMENT 39
6.1 Defining the exposure pathway 40
6.2 Quantifying each component of the exposure pathway 41
6.2.1 Source: concentration of pathogens in environmental media 41
6.2.2 Quantifying efficacy of control measures 47
6.2.3 Intake: exposure volumes and frequencies 51

7 HEALTH EFFECTS ASSESSMENT 55
7.1 Dose–response relationships 55
7.1.1 Variability in infectivity 55
7.1.2 Dose–response models 56
7.1.3 Published dose–response relationships 56
7.2 Disability-adjusted life years 57
7.3 Secondary transmission and immunity 58

8 RISK CHARACTERIZATION 61
8.1 Quantitative measures of risk 61
8.2 Probability of infection: multiple exposure events 62
8.3 Evaluate the impact of uncertainty and variability 64
8.3.1 Uncertainty analysis 64
8.3.2 Sensitivity analysis 65

9 HOW CAN QMRA SUPPORT WATER SAFETY MANAGEMENT? 71
9.1 Regulation and health-based treatment targets 71
9.2 Site-specific water and sanitation safety planning 73
9.2.1 Know your system 73
9.2.2 System assessment 74
9.2.3 Operational targets and setting critical limits 76
9.2.4 System planning and development 78
9.3 Principles of QMRA that are relevant to risk management 79

10 CONCLUSIONS AND NEXT STEPS 81

REFERENCES 83

ANNEX A CASE-STUDIES **91**

A1 Case-study 1: Pathogen risk to swimmers at non-sewage-impacted recreational beaches in the USA (Schoen & Ashbolt, 2010) 91

A2 Case-study 2: Water reclamation redesign for reducing *Cryptosporidium* risks at a recreational spray park in the USA (Weir et al., 2011) 98

A3 Case-study 3: Evaluating *Cryptosporidium* risk at a large number of drinking-water systems in France (Medema et al., 2009) 102

A4 Case-study 4: USEPA Long Term Surface Water Treatment Rule – health benefit of a new drinking-water regulation in the USA (USEPA, 2005) 106

A5 Case-study 5: Guidelines for water recycling – Setting health-based performance targets and safe use of wastewater in Australia (NWQMS, 2006) 109

A6 Case-study 6: WHO health-based criteria for evaluating household water treatment technologies 117

ANNEX B DRINKING-WATER QMRA DEMONSTRATING THE APPLICATION OF VARIABILITY AND UNCERTAINTY ANALYSIS **123**

B1 Hypothetical case-study 123

B1.1 Case description *123*

B1.2 Problem formulation *123*

B1.3 Exposure assessment *125*

B1.4 Health effects assessment *125*

B1.5 Risk characterization *125*

B2 Part A: Quantifying exposure with no local information – point estimates 125

B2.1 Data collection *125*

B2.2 Quantification of model inputs *126*

B2.3 Risk characterization *126*

B3 Part B: Quantifying exposure with some local information – point estimates 126

B3.1 Data collection *126*

B3.2 Quantification of model inputs *127*

B3.3 Risk characterization *127*

B4 Part C: Accounting for variability in exposure – point estimates 127

B4.1 Variation in source water concentration *128*

B4.2 Variation in source water concentration, treatment and consumption volume *128*

B5 Part D: Accounting for variability – stochastic simulation 130

B5.1 Data collection *130*

B5.2 Quantification of model inputs *130*

B5.3 Risk characterization *131*

B6 Part E: Accounting for parameter uncertainty – stochastic simulation 131

B6.1 Data collection *131*

B6.2 Quantification of model inputs *132*

B6.3 Risk characterization *132*

B7 Accounting for other sources of uncertainty 132

B8 References 134

ANNEX C MICROBIOLOGICAL DATA AND STATISTICAL INFERENCE **135**

C1 Microbial enumeration methods 135

C1.1 Multiple steps for concentration and purification *135*

C1.2 Characteristic of the organism that is being targeted *136*

C1.3 Quantitative or qualitative observations *137*

C2 Statistical inference 137

C2.1 Inference based on reported concentrations (Approach 1 in Fig. C.2) 138

C2.2 Inference based on original observations (Approach 2 in Fig. C.2) 139

C3 Quantifying pathogen concentration 140

C3.1 Data considerations 140

C3.2 Accounting for method recovery 141

C3.3 Accounting for viability/infectivity and human infectivity 143

C3.4 Example: modelling the distribution of concentration from pathogen data 144

C4 Faecal indicators and quantitative prediction of pathogen concentration 148

C4.1 Theoretical relationship between pathogens and FIO 149

C4.2 Implementation of pathogen to indicator ratios in QMRA 152

C4.3 Example: application of pathogen to indicator ratio in QMRA (Howard, Pedley & Tibatemwa, 2006) 153

C5 Quantifying pathogen reduction across barriers: empirical approach 156

C5.1 Data considerations 156

C5.2 Statistical analysis 157

C5.3 Example: quantifying $\log_{10}$ reduction from microbial counts 158

C5.4 Example: reduction of Giardia *across conventional drinking-water treatment* 159

C5.5 Example: reduction across a full-scale urban wetland 162

C6 Quantifying pathogen reduction across barriers: mechanistic approach 165

C6.1 Hydrologic models 165

C6.2 Pathogen inactivation 165

C7 References 166

ANNEX D DOSE–RESPONSE 171

D1 Observations 171

D1.1 Clinical studies 171

D1.2 Outbreaks 171

D2 Dose–response models 172

D2.1 Maximum risk curve: the upper limit of the single-hit model 172

D2.2 Fractional Poisson model 173

D2.3 Application of single-hit models in QMRA: adopting a conditional approach 173

D2.4 Application of single-hit models in QMRA: low-dose approximations 173

D3 Published dose–response relationships 173

D3.1 Campylobacter jejuni 174

D3.2 E. coli *O157:H7* 175

D3.3 Enteroviruses 175

D3.4 Adenovirus 176

D3.5 Rotavirus 176

D3.6 Norovirus 176

D3.7 Giardia 177

D3.8 Cryptosporidium 177

D4 References 180

ANNEX E GLOSSARY 183

ACKNOWLEDGEMENTS

The preparation of this document involved the important work of a variety of technical and scientific experts and stakeholders over several years. The document was authored by Dr Susan Petterson (Water and Health, Pty Ltd, Australia), Professor Gertjan Medema (KWR Watercycle Research Institute and Delft University of Technology, the Netherlands) and Professor Nicholas Ashbolt (University of Alberta, Canada, formerly United States Environmental Protection Agency). Dr Lesley D'Anglada (United States Environmental Protection Agency) and Dr Maggie Montgomery (World Health Organization [WHO]) additionally contributed case-studies.

The development and production of this document were supported by the following staff of the Water, Sanitation, Hygiene and Health (WSH) unit of WHO: Professor Jamie Bartram (currently University of North Carolina, United States of America (USA), Mr Robert Bos (currently International Water Association), Mr Chee Keong Chew (currently Public Utilities Board (PUB), Singapore), Ms Jennifer De France and Mr Bruce Gordon.

An international group of over 35 experts contributed to the development of this document through participation in workshops, peer review and provision of insights and text. These include:

- Rafael Bastos, Universidade Federal de Vicosa, Brazil
- Sophie Boisson, WHO, Switzerland
- Teresa Brooks, Health Canada, Canada
- Enrique Calderon, Sanitary Engineering Institute, Argentina
- Phil Callan, Consultant, Australia
- David Cunliffe, Department of Health, Australia
- Dan Deere, Water Futures, Australia
- Ana Maria de Roda Husman, National Institute for Public Health and the Environment, the Netherlands
- Takuro Endo, Ministry of Health, Labour and Welfare, Japan
- Michèle Giddings, Health Canada, Canada
- Karina Gin, Nanyang Technological University, Singapore
- Paul Hunter, University of East Anglia, United Kingdom
- H.E. Jianzhong, National University of Singapore, Singapore
- Thaís Terceiro Jorge, formerly WHO, Switzerland
- Claus Jørgensen, DHI, Denmark
- Naohiro Kishida, National Institute of Public Health, Japan
- Suresh Kumar, University of Malaya, Malaysia
- Stephanie McFadyen, Health Canada, Canada
- Rory Moses McKeown, WHO, Switzerland
- Fuminari Miura, WHO, Switzerland and University of Tokyo, Japan
- Maggie Montgomery, WHO, Switzerland
- Joanne O'Toole, Monash University, Australia
- Stephen Schaub, formerly United States Environmental Protection Agency, USA
- Jack Schijven, National Institute for Public Health and the Environment, the Netherlands
- Oliver Schmoll, WHO Regional Office for Europe, Germany
- Dai Simazaki, National Institute of Public Health, Japan
- Mark Sobsey, University of North Carolina at Chapel Hill, USA
- Jeff Soller, Soller Environmental, LLC, USA
- Thor-Axel Stenström, Swedish Institute of Communicable Disease Control, Sweden (currently Durban University of Technology, South Africa)
- Peter Teunis, National Institute for Public Health and the Environment, the Netherlands, and Rollins School of Public Health, Emory University, USA
- Andrew Trevett, WHO, Bangladesh (currently UNICEF, Kenya)

- Emily Viau, Bioenergy Frontiers, USA
- Timothy Wade, United States Environmental Protection Agency, USA
- Dano Wilusz, United States Department of State, USA

Ms Marla Sheffer of Ottawa, Canada, was responsible for editing this document, and Ms Penny Ward and Mr Eben Johnson provided secretarial and administrative support throughout the document development process and to individual meetings and workshops.

WHO gratefully acknowledges the financial and technical support provided by the Environmental Protection Agency and the Department of State, USA; the Department for International Development, United Kingdom; the Ministry of Environment and Water Resources, Singapore; and the Ministry of Health, Labour and Welfare, Japan.

ABBREVIATIONS

ABS	Australian Bureau of Statistics
AIC	Akaike information criterion
AIDS	acquired immunodeficiency syndrome
CDF	cumulative density function
cDNA	complementary deoxyribonucleic acid
CFU	colony-forming unit
CIRSEE	Centre International de Recherche sur l'Eau et l'Environnement
ct	concentration of disinfectant × time
DALY	disability-adjusted life year
DAPI	diaminophenylindole
DNA	deoxyribonucleic acid
FAO	Food and Agriculture Organization of the United Nations
FFU	focus-forming unit
FIO	faecal indicator organisms
FS	factor sensitivity
GDWQ	WHO Guidelines for drinking-water quality
GREC	WHO Guidelines for safe recreational water environments
GWEG	WHO Guidelines for the safe use of wastewater, excreta and greywater
HIV	human immunodeficiency virus
HWT	household water treatment
HWTS	household water treatment and safe storage
ICR	Information Collection Rule
ICRSSL	Information Collection Rule Supplemental Surveys Large
ICRSSM	Information Collection Rule Supplemental Surveys Medium
ID_{50}	50% infectious dose
IFA	immunofluorescence assay
IWA	International Water Association
LR	log_{10} reduction
LT2ESWTR	Long Term 2 Enhanced Surface Water Treatment Rule
MCMC	Markov chain Monte Carlo
MEC	mean elimination capacity
MLE	maximum likelihood estimator
MPN	most probable number
MPNCU	most probable number of cytopathic units
MST	microbial source tracking
NRSA	nominal range sensitivity analysis
NTU	nephelometric turbidity unit
PCR	polymerase chain reaction
PDF	probability density function
PDU	PCR detectable unit
PFU	plaque-forming unit
pppy	per person per year
PUB	Public Utilities Board
QMRA	quantitative microbial risk assessment

qPCR	quantitative polymerase chain reaction
RIVM	National Institute of Public Health and the Environment (the Netherlands)
RNA	ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
SC	step characteristic
SRC	sulfite-reducing clostridia
SSP	sanitation safety plan
$TCID_{50}$	50% tissue culture infective dose
TNTC	too numerous to count
UNICEF	United Nations Children's Fund
USA	United States of America
USEPA	United States Environmental Protection Agency
UV	ultraviolet
VBNC	viable but non-culturable
WHO	World Health Organization
WSP	water safety plan
YLD	years living with a disability
YLL	years of life lost

EXECUTIVE SUMMARY

The World Health Organization's (WHO) water quality guidelines recommend a preventive, risk-based approach to water quality management from source to exposure for the management of microbial hazards. Risk assessment therefore plays a central role in implementing the guidelines through the development of water safety plans (WSPs) and sanitation safety plans (SSPs). Several risk assessment approaches are available, ranging from risk scoring in sanitary inspections and risk matrices to quantitative microbial risk assessment (QMRA). Guidance is available for sanitary inspections[1] and for the use of risk matrices in WSPs[2,3] and SSPs.[4] However, although much academic literature is available on QMRA, limited documentation exists to facilitate the consistent application of QMRA in the practice of water supply, water reuse or water recreation.

The purpose of this document is therefore threefold:

1) to present a harmonized framework for the application of QMRA to evaluate risks associated with faecal pathogens for the drinking-water, wastewater and recreational water pathways;
2) to provide guidance on the interpretation of scientific data (pathogen concentrations in different water matrices, data on pathogen removal by different barriers under different conditions, data on exposure volumes and dose–response relationships) for application in QMRA; and
3) to provide guidance on how to set up and interpret the results of QMRA in order for it to be an effective support for water safety management.

The intended audience for this document is regulators, water supply and sanitation system engineers and managers, as well as scientists working to ensure that water-related health risks are minimized in a particular setting. The document serves to support an understanding of how QMRA works, the value of QMRA for water safety management and considerations in applying QMRA. By capturing QMRA methods and application in one document, the intention is to achieve crossover and understanding for readers from different backgrounds.

The four steps of the harmonized QMRA framework are:

1) problem formulation;
2) exposure assessment;
3) health effects assessment; and
4) risk characterization.

Each step of the QMRA is presented in detail within the document. Briefly, in *problem formulation*, the scope and purpose of the risk assessment are carefully defined in consultation with risk managers. In order for the QMRA to be useful, it is critical that there is a transparent dialogue between the risk manager and risk assessor throughout the assessment process. This begins at the scoping, planning and problem formulation stage to ensure that the risk assessment addresses the needs of the risk manager. It is essential that the assessment is defined in such a way to ensure that the calculations (including uncertainty analysis) are targeted towards the objectives and needs of the risk manager.

Relying on the scope defined during problem formulation, the *exposure assessment* is undertaken to quantify the magnitude and frequency of exposure to reference pathogens via the defined exposure pathways and hazardous events. The exposure assessment involves three steps. Firstly, the exposure pathway is defined in terms of pathogen sources (the point of initial pathogen quantification), barriers or control measures, and mechanisms of exposure. Secondly, each component of the exposure pathway is quantified based on the best available scientific data and an understanding of the expected variability and uncertainty associated with the model variables (Chapter 6 and Annex C contain detailed discussion and resources to support this quantification). Thirdly and finally, the magnitude and frequency of exposure are quantified for the range of defined scenarios.

[1] Guidelines for drinking-water quality, second edition. Vol. 3. Surveillance and control of community supplies. Geneva: World Health Organization; 1997.

[2] Water safety plan manual. Step-by-step risk management for drinking-water suppliers. Geneva: World Health Organization; 2009.

[3] Water safety planning for small community water supplies. Step-by-step risk management guidance for drinking-water supplies in small communities. Geneva: World Health Organization; 2012.

[4] Sanitation safety planning: manual for safe use and disposal of wastewater, greywater and excreta. Geneva: World Health Organization; 2015.

Similarly relying on the scope defined during problem formulation, the *health effects* data are compiled for each reference pathogen. This will include a dose–response relationship and any subsequent health outcomes, sequelae or disease burden estimates. Appropriate selection of these values to represent the pathogen of interest and the exposed population is important to the outcome of the risk assessment. More detailed discussion of these factors and supporting information for dose–response is included in Chapter 7 and Annex D.

Finally, the exposure and health effects information is combined by calculating the risk for the defined conditions and scenarios during *risk characterization*. The outcome of the risk characterization must feed directly back to the risk management questions defined during problem formulation. Uncertainty analysis and/or sensitivity analysis can be extremely valuable for testing the robustness of the risk calculations to support the risk management decision.

In undertaking a QMRA, the following overarching principles need to be considered so that the QMRA is useful for risk management:

- The QMRA is driven by the risk management need; the decision (and the required level of certainty for that decision) has driven the problem formulation.
- The QMRA is fit-for-purpose in terms of scope and level of detail.
- The QMRA is based on exposure scenarios consistent with the risk management options, local context and data constraints.
- Factors influencing variability associated with each component of the exposure are considered and documented.
- The data and information used to quantify exposure are the most appropriate for use and weighed against the information in the scientific literature.
- The representativeness of published dose–response models is evaluated.
- The QMRA is focused on the population exposed. As appropriate, it will give consideration to susceptible population subgroups and life stages.
- The quantities of risk calculated are driven by the problem and directly relate to the risk management decision.
- The influence of uncertainties on the outcome of the assessment is evaluated.

These principles, combined with documentation, model verification and review, can result in a transparent assessment that enables peers, risk managers and other stakeholders to understand the basis for model development, data selection and processing, and assumptions. This will help to increase confidence that the QMRA approach taken is adequate for the risk management questions that it addresses and that the conclusions about risk management options are credible.

Using this framework, QMRA can be of great value for water safety management at both the regulatory and site-specific levels. Many examples of the application of QMRA for water safety management in different settings and for different purposes are given in Chapter 9 and the case-studies in Annex A. Specific applications to support WSPs and SSPs include determining whether microbial health-based targets are achieved, evaluating monitoring requirements and establishing critical limits to support system planning and development, choosing between potential interventions and cost–benefit analysis.

1 INTRODUCTION: THE EVIDENCE BASE FOR ADEQUATE RISK MANAGEMENT

1.1 Moving towards preventive, risk-based water safety management

According to the most recent World Health Organization (WHO) estimates, 842 000 diarrhoeal deaths in low- and middle-income countries are caused by inadequate drinking-water, sanitation and handwashing practices (WHO, 2014a). Although this number represents a large drop in water, sanitation and hygiene–related diarrhoeal disease in the last decade, the burden is still unnecessarily large. In addition, outbreaks of waterborne disease continue to occur in both developed and developing countries, leading to loss of life and significant disease and economic burden for individuals and communities (e.g. Craun et al., 2010; Collier et al., 2012).

These findings highlight the importance of ensuring the microbial safety of water and sanitation services. The WHO water quality guidelines therefore emphasize that primary attention should be given to managing microbial hazards in the water cycle. In controlling these, there has traditionally been great emphasis on the examination of faecal indicator bacteria. At the turn of the millennium, it was increasingly recognized that this basis for the risk management of microbial hazards was:

- too little – viruses and parasites are significant microbial hazards in water; their fate and transport in the environment and water treatment processes are very different from those of (faecal indicator) bacteria; and outbreaks of waterborne disease have occurred while the water complied with the guidelines for faecal indicator bacteria; and
- too late – by the time the faecal indicator bacteria examination highlighted a potential health problem, the exposure to the water had already occurred.

For these reasons, the WHO water quality guidelines recommend a preventive, risk-based approach to water quality management, covering source to exposure. The risk-based approach was first adopted by WHO in 1999 as the Stockholm Framework. The framework provides a common conceptual approach to assessing water quality hazards and managing associated risks (Fig. 1.1). Although it was developed for the control of waterborne infectious diseases, it also serves as a basis for the control of toxic chemicals and other health hazards. This framework involves a systematic assessment of risks, definition of health-based risk management objectives and planning the appropriate risk management response (Fig. 1.1; see also Fewtrell & Bartram, 2001; WHO, in preparation). The systematic risk assessment advocated in this framework provides an evidence base to guide appropriate risk management decisions. The WHO water-related guidelines have been structured in line with the Stockholm Framework.

The instruments to make the framework and guidelines operational are preventive risk management plans for drinking-water, recreational water and sanitation. WHO has published step-by-step guidance documents for developing and implementing water safety plans (WSPs) for piped water supply systems (Bartram et al., 2009) and for water supply systems for small communities (WHO, 2012, 2014b), as well as for developing and implementing

sanitation safety plans (SSPs) (WHO, 2015a). Fig. 1.1 provides links to each WHO water quality guideline and their associated water safety planning resources.

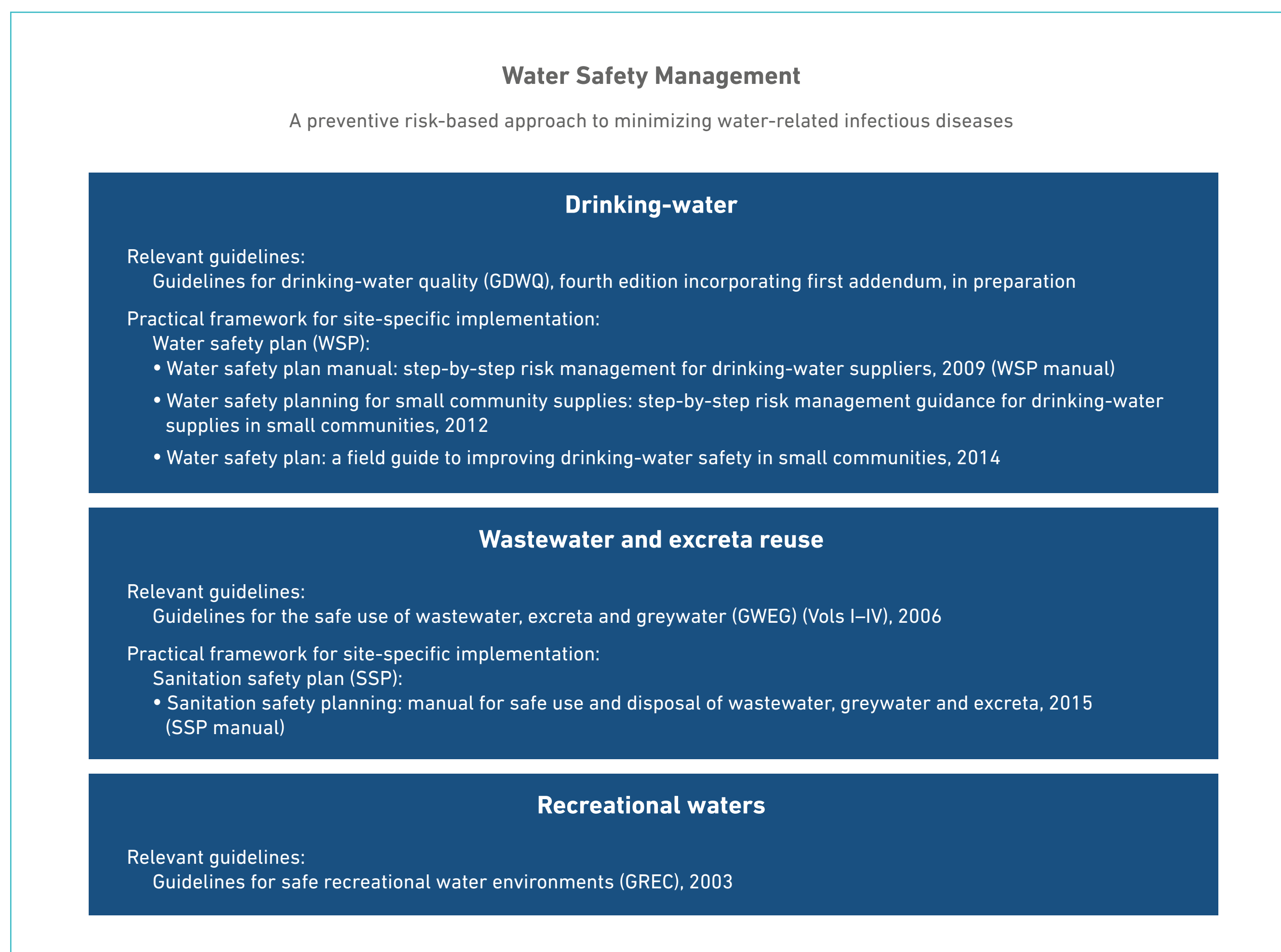


Fig. 1.1 Water safety management in the WHO water quality guidelines

1.2 Purpose of this document

Risk assessment plays a central role in the WHO water quality guidelines and in WSPs and SSPs. Several risk assessment approaches are available, ranging from risk scoring in sanitary inspections through use of sanitary inspection forms to risk matrices, which are often the risk assessment approach used in WSPs and SSPs, to quantitative microbial risk assessment (QMRA) (see Box 1.1). Guidance is available for sanitary inspections (WHO, 1997, 2014b) and for the use of risk matrices in WSPs (Bartram et al., 2009) and SSPs (WHO, 2015a). However, although much academic literature is available on QMRA, limited documentation exists to facilitate the consistent application of QMRA in the practice of water supply, water reuse or water recreation.

The purpose of the document is therefore threefold:

1) to present a harmonized framework for the application of QMRA to evaluate risks associated with faecal pathogens for the drinking-water, wastewater and recreational water pathways;
2) to provide guidance on the interpretation of scientific evidence (pathogen concentrations in different water matrices, data on pathogen removal by different barriers under different conditions, data on exposure volumes and dose–response relationships) for application in QMRA; and
3) to provide guidance on how to set up and interpret the results of QMRA in order for it to be an effective support for water safety management.

Box 1.1 Epidemiology and risk assessment

While risk scoring, risk matrices and QMRA are approaches that assess the *potential* risk to human health, epidemiology is the science that attempts to measure the *actual* distribution and determinants of health events. Epidemiology has particular value in estimating the population incidence of certain diseases and determining the proportion of such cases that are attributable to different exposures. As such, epidemiological studies can provide estimates of the actual disease burden due to various exposures, including disease associated with water exposures (see Blumental et al., 2001; Hunter et al., 2003). However, epidemiological studies are often expensive (requiring the recruitment of large numbers of participants) or time consuming (necessitating follow-up of study participants for months), or both. Therefore, although epidemiological studies provide an essential basis for understanding the actual health outcomes and burden of disease associated with water, such studies are often not applicable to individual local supplies and settings, because the population size is not big enough, because there are too few cases to give sufficient statistical power or because of the cost and effort needed to undertake them. In these contexts, QMRA and other risk assessment techniques are more feasible and may in fact be the only approach open to water supply managers to assess any risk to their consumers. Consequently, this document focuses on the complementary risk assessment approaches that are available for source-to-exposure water safety management.

1.3 Intended audience and reading guide

The intended audience is regulators, water supply and sanitation system engineers and managers, as well as scientists working to ensure that water-related health risks are minimized in a particular setting. The document serves to support an understanding of how QMRA works, the value of QMRA for water safety management and considerations in applying QMRA. Readers who are primarily interested in understanding how QMRA fits within the risk assessment spectrum, the basics of QMRA and learning how QMRA can support decision-making will find this information mostly in Chapters 1, 2 and 9 and in the case-studies in Annex A. Readers who are interested in gaining an in-depth understanding of the QMRA process and considerations for undertaking a QMRA (e.g. how to work with microbiological data on pathogens, indicator organisms, barrier efficacy, dose–response data, probabilistic models) will find this information in Chapters 3–8 and Annexes B–D. By capturing QMRA methods and application in one document, the intention is to achieve crossover and understanding for readers from these different backgrounds.

1.4 Structure of this document

Introduction (Chapter 1): In Chapter 1, the importance and key elements of preventive, risk-based water safety management are described.

Spectrum of risk assessment approaches (Chapter 2): This chapter gives an overview of risk assessment approaches that are in use in water safety management, how they provide an evidence base for risk management decisions and their strengths and limitations.

Framework for water-related QMRA (Chapter 3): The four-step process (problem formulation, exposure assessment, health effects assessment and risk characterization) is described to facilitate the harmonized implementation of QMRA for drinking-water, wastewater and recreational exposures. Initial planning considerations for conducting QMRA are highlighted, including team characteristics, tools required and how to work in resource-limited settings. The principles of a QMRA that is relevant for risk management are given.

Variability and uncertainty (Chapter 4): The concepts of variability and uncertainty as they relate to the quantification of waterborne infectious risks are introduced. Appropriate consideration of these factors is critical to the overall implementation of the QMRA framework and to ensure that the risk assessment results are useful to support decision-making.

Chapters 5–8 expand on the content of each of the four steps in the QMRA framework:

1) **Problem formulation (Chapter 5):** Including how to define the scope and purpose of the QMRA so that it will meet the objectives of the study and the needs of risk managers.
2) **Exposure assessment (Chapter 6):** The process of simplifying a complex environmental system into a defined exposure pathway that will meet the objectives of the risk assessment.

3) **Health effects assessment**[1] **(Chapter 7):** Considerations for evaluating and quantifying the health effects (short and long term) associated with exposure to faecal pathogens.
4) **Risk characterization (Chapter 8):** Considerations for combining the content of steps 1–3 into quantitative measures of risk that are meaningful for addressing the defined problem.

How can QMRA support water safety management? (Chapter 9): The role of QMRA in supporting the management of infectious disease risks is discussed, and examples are presented.

Conclusions and next steps (Chapter 10): Overview of further opportunities for development of QMRA as a tool to support risk managers.

Annex A: Case-studies: Risk management is the process of evaluating alternative options and selecting the most appropriate interventions. Annexed to this document is a collection of case-studies (Table 1.1) that demonstrate the practical application of QMRA for different water safety management strategies. Studies were selected to represent a range of QMRA approaches, different contexts, including drinking-water, wastewater and recreational water, and different management objectives (e.g. household, utility, national levels).

Table 1.1 Overview of case-studies

Risk management objective		Case-study	Reference	Related WHO guideline
Evaluation of hazards	1	Pathogen risk to swimmers at non-sewage-impacted recreational beaches in the USA	Schoen & Ashbolt (2010)	GREC
Evaluate alternative options	2	Water reclamation redesign for reducing *Cryptosporidium* risks at a recreational spray park in the USA	Weir et al. (2011)	GREC
Determine priorities	3	Evaluating *Cryptosporidium* risk at a large number of drinking-water systems in France	Medema et al. (2009)	GDWQ
Cost–benefit	4	USEPA Long Term Surface Water Treatment Rule – health benefit of new drinking-water regulation in the USA	USEPA (2005)	GDWQ
Setting health-based performance targets	5	Guidelines for water recycling – Setting health-based performance targets and safe use of wastewater in Australia	NWQMS (2006)	GWEG
	6	WHO health-based criteria for evaluating household water treatment technologies	WHO (2011)	GDWQ

USEPA: United States Environmental Protection Agency

Annex B: Drinking-water QMRA demonstrating the application of variability and uncertainty analysis: This annex illustrates the concepts presented in Chapter 4 and provides a hypothetical worked example of QMRA of *Cryptosporidium* in drinking-water. The example begins simply and progresses to a more complex analysis, including variability and uncertainty assessment.

Annex C: Microbiological data and statistical inference: This annex is targeted at water management teams that have limited knowledge of microbial laboratory and data analysis techniques. It provides practical guidance and examples for performing basic to more complex assays and for analysing microbial data.

Annex D: Dose–response: A more detailed overview of published models for each of the key reference pathogens is given.

Annex E: Glossary: Definitions of selected terms used throughout this document are included.

[1] Referred to as hazard characterization in some QMRA frameworks.

2 | SPECTRUM OF RISK ASSESSMENT APPROACHES

In this chapter, the concept of risk assessment is explained, and the various approaches ranging from sanitary inspection to QMRA are introduced.

Risk assessment is a key element of the Stockholm Framework and is used in all water-related WHO guidelines. This chapter introduces the spectrum of risk assessment approaches that are in use in water safety management, ranging from sanitary inspections to QMRA, in order to provide an overview of the application, information requirements, strengths and limitations of each approach. All text and examples in this chapter are on drinking-water systems to avoid repetition and to facilitate a more direct comparison of approaches. The basic principles are also applicable to recreational water or water reuse.

For drinking-water systems, risk assessment is an integral part of developing and implementing WSPs. The purpose of the risk assessment is to identify and evaluate the health risks associated with the water supply, to determine if the health hazards are adequately controlled, to inform operation and management of the water supply and to identify necessary improvements and upgrades to ensure the delivery of safe drinking-water. Health hazards can be pathogenic microorganisms, chemicals or radioactive substances in concentrations that have an adverse health impact. The process of risk assessment is a systematic evaluation of:

1) hazards – in this document, pathogenic microorganisms (pathogens) that may have an adverse impact on the health of the people who drink the water;
2) hazardous events – events that may introduce pathogens into the water supply or fail to remove them. These events may occur at every step of the water supply chain – for example, at the source (e.g. rain events that flush human or animal faecal waste into the water supply), in treatment (e.g. failures in filtration or disinfection), in the distribution network (e.g. improper repair work introducing microbial contamination) and in households (e.g. handling storage containers with dirty hands);
3) the adequacy of the controls to prevent contamination – control measures that are or could be put in place to prevent these hazards from occurring, to remove these hazards from the water supply system or to reduce these hazards to an acceptable level. These can be engineered controls, such as a water treatment process, as well as non-engineered measures, such as hygiene protocols for repair works on the water distribution network.

Hazard: a biological (or physical, chemical, radiological) agent that can cause harm to public health.

Hazardous event: an event that introduces a hazard into the water supply or fails to remove it.

Risk: an assessment of the likelihood that a hazardous event occurs and the severity or consequence of the hazard.

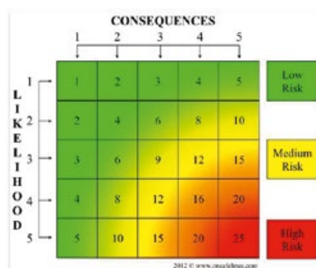
There are potentially many hazards, many hazardous events and many control measures. The objective of risk assessment is to identify those risks that are critical for the safety of a specific water supply system and to help to select the best steps to improve the safety of the system. Hence, to help risk management focus on the needed improvements that contribute most to water supply safety, risk assessment needs some form of classification or quantification of the risks in terms of health impact.

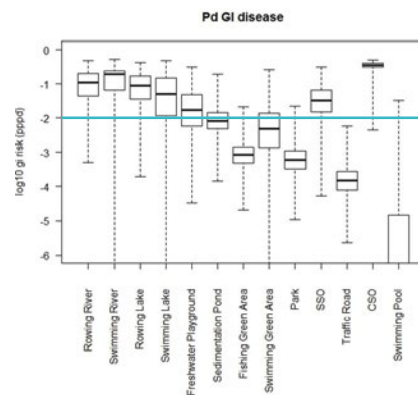
The outcome of the risk assessment process is an evaluation of the safety of the water supply system, based on an understanding of the hazards, hazardous events and validity of control measures in the water supply system and their relative significance. Through the risk assessment process, the multidisciplinary team that is preparing the WSP will understand the hazards and hazardous events that may have an impact on the water supply system and the adequacy of the barriers and control measures that are in place. The risk assessment may show that the system is safe and provide the justification for this conclusion, or the risk assessment process may highlight shortcomings in the current water supply system design, operation or maintenance and support the prioritization of these shortcomings according to their health risk and thus their priority for management. This provides the evidence-based justification for an improvement plan for the water supply system, focusing on the priorities highlighted by the risk assessment process. Risk assessment is therefore a decision support tool that provides the risk manager with an objective and rational picture of what is known (or assumed, based on expert judgement) about the risks associated with the water supply. It is imperative that risk assessment is quantified in some manner; otherwise, the risk manager is faced with all possible risks without an indication of their significance.

A range of approaches to conduct a risk assessment is available:

- **Sanitary inspection:** An on-site visual evaluation of observable features and conditions at or in the vicinity of the water supply that may lead to an unsafe supply. Sanitary inspections are typically based on standardized forms/checklists to identify the most common issues that may lead to the introduction of hazards into a system. This approach has been developed and promoted as a simple and effective tool for small water supplies (WHO, 1997) and, more recently, as part of WSPs for small supplies (WHO, 2012b, 2014b). At the local level, sanitary inspections will assist operators, as well as water and health officers, in the identification of the most important causes and pathways of contamination and control options to prevent or minimize contamination. The approach is also useful when applied more broadly to inform regional or national priorities for improving small supplies.
- **Risk matrix:** The risk assessment approach that makes a qualitative or semiquantitative evaluation of the likelihood that a hazardous event will occur and the severity or consequence of the hazard and combines them into a risk score or risk rating. The approach relies on expert judgement and can be undertaken at different levels of detail. It has been applied as a simple, common approach to evaluate the range of different (water quality–associated) risks in the WSP.
- **QMRA:** A formal, quantitative risk assessment approach that combines scientific knowledge about the presence and nature of pathogens, their potential fate and transport in the water cycle, the routes of exposure of humans and the health effects that may result from this exposure, as well as the effect of natural and engineered barriers and hygiene measures. All this knowledge is combined into a single assessment that allows evidence-based, proportionate, transparent and coherent management of the risk of waterborne infectious disease transmission. QMRA has developed as a scientific discipline over the last two decades and has been embedded in the WHO water-related guidelines (WHO, 2003, 2006a,b, in preparation).

All aforementioned risk assessment approaches are valid – their use is context specific and will depend on, among others, human resources (personnel, skills, access to support institutions) and type of supply system (small community-managed supplies versus larger utility-managed supplies). They represent a continuum from more simple to more detailed and from more expert judgement to more evidence-based assessment of the risks (Fig. 2.1). In general, risk assessments should be as simple as possible, finding the right balance between more detail and evidence base and the use of assumptions and expert judgement. The right balance is the one that is considered adequate to inform on risk management options. Where simplified risk assessments are sufficient to aid risk management decisions, there is no need to undertake more complicated assessments. The approaches and their strengths and limitations in supporting adequate water safety management are given in the following overview of the scale of approaches for (microbial) risk assessment and the increased level of detail in input and output of the different approaches.





SANITARY INSPECTION

RISK MATRIX

QUANTITATIVE MICROBIAL RISK ASSESSMENT
Screening In depth

LEVEL OF KNOWLEDGE AND RESOURCES
LEVEL OF DETAIL IN REQUIRED INFORMATION
UNDERSTANDING OF HAZARDS & CONTROLS
LEVEL OF EVIDENCE BASE IN RISK ASSESSMENT

Fig. 2.1 The scale of microbial risk assessment: more detailed input facilitates more evidence-based risk management

2.1 Sanitary inspection

Sanitary inspection is a powerful and generally applicable tool for the risk assessment of water supply systems. It is widely used in small water supply settings to support the identification and management of high-priority risk factors. Sanitary inspections can also support WSP implementation, including the identification of hazardous events and potential control measures. Results of sanitary inspections can inform more systematic risk assessments that may be conducted within a WSP (e.g. the risk matrix).

Community and/or health and water officers with knowledge about water supply and health hazards visit the water supply system and inspect the presence of observable contamination sources and conditions that may lead to the introduction of hazards into the water system, the potential occurrence of hazardous events and how these are controlled by control measures. Sanitary inspections are frequently used for point sources (e.g. wells, springs), focusing on the close vicinity to the source and the integrity and conditions of abstraction or storage infrastructures. Sanitary inspections can also cover storage tanks/service reservoirs and, to a limited extent, distribution and treatment systems.

Sanitary inspections typically make use of standardized “sanitary inspection forms” containing a systematic checklist of a limited number of specific questions. These checklists address the most basic and common factors that may lead to contamination of the water system. Sanitary inspection forms for a range of small community and private supplies can be found in Volume 3 of the second edition of the WHO GDWQ (WHO, 1997) and the WHO WSP field guide for small communities (WHO, 2014b). An example of a checklist for a spring source is given in Box 2.1.

Box 2.1 Sanitary inspection checklist for spring sources

SANITARY INSPECTION FORM 4
SPRING SOURCE

I. General information

a. Name of village or town: ....

b. Location and/or name of spring: ....

c. Date of inspection: ....

d. Weather conditions during inspection: ....

Note. *If there is more than one spring source in your community, or if the community uses other water sources (such as dug wells or boreholes), carry out sanitary inspections for these sources too.*

Note. *If the spring serves a storage reservoir or directly feeds into a piped distribution system, also carry out sanitary inspections using the forms "Storage reservoirs" and/or "Public/yard taps and piped distribution" respectively.*

Note. *If consumers store water in homes, also regularly inspect water storage and handling in homes using the sanitary inspection form "Collection and household containers".*

II. Specific questions for assessment

1. Is the spring box absent or faulty? Yes | No
2. Is the brick wall or backfill area protecting the spring faulty or eroded? Yes | No
3. If there is a spring box, is the inspection cover absent, faulty or unsanitary, or is the concrete around the cover damaged? Yes | No
4. Does spilt water flood the collection area? Yes | No
5. Is the spring box unfenced, or is the fence inadequate or faulty? Yes | No
6. Can animals have access within 10 metres of the spring? Yes | No
7. Is there a latrine uphill and/or within 30 metres of the spring? Yes | No
8. Does surface water collect uphill within 30 metres of the spring? Yes | No
9. Is the diversion ditch above the spring absent or non-functional? Yes | No
10. Are there any other sources of pollution uphill of the spring (such as animal breeding, cultivation, roads, garages, craft enterprises or waste)? Yes | No

Total score of risk factors as total number of "YES" answers:...

III. Results and comments

a. Sanitary inspection risk score (tick appropriate box):

☐ *Very high risk* *Risk score: 9–10* ☐ *High risk* *Risk score: 6–8* ☐ *Medium risk* *Risk score: 3–5* ☐ *Low risk* *Risk score: 0–2*

b. Important points of risk noted and reported on the reverse of this form:
- list according to question numbers 1–10
- additional comments

IV. Names and signatures of assessors:..

Source: WHO (2014b)

Sanitary inspection forms should be completed using a mixture of on-site inspection and interviewing of community operators or members. These forms are generally structured in such a way that the questions can be answered only by YES or NO. The YES answer indicates the presence of a risk factor, whereas the NO answer indicates the absence of this factor. Each YES answer should trigger follow-up actions at the local level to address the identified risk factors (e.g. by rehabilitation and/or putting in place additional control measures).

The checklists also provide a simple quantitative classification of the level of safety of the water supply system (e.g. very high risk, high risk, medium risk and low risk), by counting the number of YES answers. Such a risk scoring system is particularly useful when sanitary inspection forms are used more broadly – for example, as part of a surveillance programme. It can help to determine the status of small water supply systems and inform regional and national priorities. For instance, results can shed light on which systems are the most "risky" (e.g. based on supply type or location) and which risk factors fail the most frequently (see Table 2.1).

Table 2.1 Frequency of risk factors occurring in protected springs, identified by sanitary inspection in Ethiopia

	Questions for the sanitary risk inspection	Risk frequency (%)
PROTECTED SPRING: 316 sites inspected		
1	Is the collection/spring box absent or faulty?	38.3
2	Is the masonry protecting the spring absent or faulty?	33.9
3	Is the backfill area behind the retaining wall absent or eroded?	34.8
4	Does spilled water flood the collection area?	47.2
5	Is the fence absent or faulty?	76.6
6	Can animals have access within 10 m of the spring?	94.6
7	Is there a latrine uphill and/or within 30 m of the spring?	9.5
8	Does surface water collect uphill of the spring?	18.0
9	Is the diversion ditch above the spring absent or non-functional?	87.0
10	Are there any other sources of pollution uphill of the spring (e.g. solid waste)?	26.9

Source: Tadesse et al. (2010)

It is valuable to complement the sanitary inspection with water quality monitoring for key microbial parameters (Box 2.2). The combination of water quality assessment and sanitary inspection was suggested by Lloyd & Bartram (1991), was recommended in Volume 3 of the second edition of the WHO GDWQ (WHO, 1997) and is embedded in the current edition of the WHO GDWQ (WHO, in preparation).

Box 2.2 Complementing sanitary inspection with microbial water quality monitoring

Water quality monitoring can pick up contamination events due to risks that were not easily identified or were overlooked in a sanitary inspection, and, vice versa, sanitary inspection may identify contamination sources and pathways that are not picked up by monitoring. Further, where water quality monitoring results are available, these should feed into the sanitary inspection process. The consideration of sanitary inspection combined with microbiological monitoring provides a useful dialogue between water supply and water quality experts. Together, they confirm the validity of each other's findings and support identification of the risk factors that are most likely to lead to health risks. In contrast, if limited resources are available to conduct water quality monitoring, sanitary inspection alone can provide valuable information to target management actions.

Sanitary inspections with microbiological monitoring can be used as a tool to evaluate individual water systems and inform the local management about risks and controls and suggest improvements as well as to inform regional or national priorities, to map the status of the water supply systems and inform regional or national authorities about priorities for risk management (as shown in Table 2.2).

Table 2.2 Combining sanitary inspection and water quality monitoring in a national survey of water supply system safety: prioritizing surveillance efforts and interventions

E. coli count	Sanitary inspection risk score			
	0–2	3–5	6–8	9–10
>100	17	57	31	4
11–100	45	122	55	5
1–10	26	58	22	3
0	426	609	114	3

Low risk: no action required	Intermediate risk: low action priority	High risk: higher action priority	Very high risk: urgent action required
27%	40%	13%	20%

Source: adapted from Tadesse et al. (2010)

2.1.1 Strengths

Sanitary inspections are simple and require few resources. A sanitary inspection requires a qualified individual to visit the water supply site and thoroughly review the local situation for conditions that may result in contamination. Sanitary inspections are typically based on standardized forms or checklists with simple and clear questions about the most important sources, conditions and deficiencies that may lead to hazards entering the water system. They have been used successfully in many different countries to evaluate water supply systems for small supplies. Sanitary inspections are a powerful on-site tool to enhance the knowledge about the water supply system (technical, operations, local conditions and practices), identify potential sources and pathways of contamination and thus point to required improvements and additional controls. When repeated, the sanitary inspection may also identify changes in risk that occur over time and evaluate the impact of improvement policies. Results from sanitary inspection forms are useful at an individual supply level as well as when applied as part of a surveillance programme, to inform regional and national priorities. At the regional and national levels, sanitary inspection scores can also be combined with microbial monitoring results, particularly faecal indicator bacteria, therefore further informing regional and national priorities.

2.1.2 Limitations

Sanitary inspections rely on a limited number of specific questions and therefore do not capture all hazards and hazardous events. In addition, a single sanitary inspection will not capture the variability in the conditions and practices that occur over time. Time can be incorporated by interviewing community members and repeated inspections. The risk factors in the recommended checklists are a necessary simplification and have been based on expert judgement. Hence, it is important to modify sanitary inspection forms, including risk factors, to reflect the local context. Similarly, the equal weight of the risk factors in the risk score is a simplification; therefore, sanitary inspection form results cannot be used to prioritize risk factors.

Although sanitary inspection results can be combined with microbial water quality monitoring results, particularly faecal indicator bacteria, limitations with the use of *Escherichia coli* or thermotolerant coliforms as indicators for the potential presence of faecal pathogens limit the usefulness of this approach. Compared with enteric viruses and parasitic protozoa, the faecal indicator bacteria are more sensitive to chlorine and survive for a shorter period in the water environment; compared with enteric viruses, they are more readily removed by filtration processes and soil passage. Accordingly, the *presence* of *E. coli* (or thermotolerant coliforms) is a better indicator of the potential *presence* of enteric pathogens than the *absence* of *E. coli* is for the *absence* of enteric pathogens.

2.2 Risk matrix

The WSP guidance manual (Bartram et al. 2009) describes the four steps of the system assessment:

1) Describe the water supply system from catchment to consumer in sufficient detail to allow for steps 2–4.
2) Identify the hazards and hazardous events and assess the risks for each pair of hazard and hazardous event – that is, the likelihood of each hazardous event occurring and the severity of each corresponding hazard.
3) Determine and validate control measures, in terms of both long-term performance and potential failures, to evaluate how well hazards are under control, and reassess and prioritize the residual risks.
4) Develop, implement and maintain an improvement and upgrade plan for significant risks where control is absent or considered inadequate.

The risk assessment is conducted by the WSP team and is a systematic evaluation of the risk posed by each (relevant) hazard and hazardous event, from catchment to consumer. To determine the magnitude of the risks and to prioritize the risks, a risk matrix is recommended (Bartram et al., 2009), in which the likelihood of a hazardous event occurring and the severity or consequence of the hazard are rated separately and subsequently combined into an overall risk score (Fig. 2.2). In its most basic form, this can be a qualitative assessment based on expert judgement ranking risks as, for example, significant, uncertain or insignificant (Bartram et al., 2009; WHO, 2012b). A more extensive approach is a semiquantitative assessment in which a risk matrix is used to derive a final risk score. The semiquantitative approach also relies on expert judgement and was developed as a simple, common approach to evaluate the range of different (water quality–associated) risks in the WSP (Deere et al., 2001; Bartram et al., 2009) and is used in many WSPs worldwide.

		Severity or consequence				
		Insignificant or no impact - Rating: 1	Minor compliance impact - Rating: 2	Moderate aesthetic impact - Rating: 3	Major regulatory impact - Rating: 4	Catastrophic public health impact - Rating: 5
Likelihood or frequency	Almost certain / Once a day - Rating: 5	5	10	15	20	25
	Likely / Once a week - Rating: 4	4	8	12	16	20
	Moderate / Once a month - Rating: 3	3	6	9	12	15
	Unlikely / Once a year - Rating: 2	2	4	6	8	10
	Rare / Once every 5 years - Rating: 1	1	2	3	4	5

Risk score	<6	6–9	10–15	>15
Risk rating	Low	Medium	High	Very high

Fig. 2.2 Semiquantitative risk matrix (Bartram et al., 2009)

Table 2.3 gives some examples of the semiquantitative approach. Risk matrices can come with different levels of complexity in terms of the number of categories for assessing the likelihood of a hazard event occurrence and the severity of the hazard; 3 × 3 and 5 × 5 risk matrices are commonly used in practice. In using a risk matrix, clear definitions for each of the likelihood and severity categories (e.g. likely, moderate) need to be drawn up and agreed upon before undertaking the risk assessment. This should aid in consistency in the assessment for all parts of the water supply system and over time. Similarly, definitions for what is meant by the risk categories (e.g. very high and high) should be determined to clearly flag the urgency of required attention and to facilitate action.

There are many potential hazards and hazardous events, and this may make risk assessment a labour-intensive exercise. However, many hazards and hazardous events are not unique for each water supply system. To facilitate the risk assessment process, typical hazards and hazardous events and/or control measures have been collected in checklists, databases and tools (e.g. Beuken et al., 2008; Ministry of Health, 2014). Some of these and additional tools to support risk assessment are available on the WSPortal (www.wsportal.org), an International Water Association (IWA)/WHO website on WSP resources. There are also some examples of capturing the general knowledge on hazards and hazardous events in predefined severity scores to facilitate risk assessment, particularly for small or private supplies (Scottish Executive, 2006; Environmental Protection Agency, 2010).

Table 2.3 Example of risk assessment elements using the risk matrix for microbial hazards in a surface water supply

Hazardous event	Likelihood	Severity	Score	Risk	Justification
High pathogen concentration at water intake during heavy rainfall	3	4	12	High	Observation of overflowing sewer manholes frequently reported by local inhabitants upstream of the intake following heavy rain; observation confirmed during site inspections; following heavy rainfall, increased levels of *E. coli* have been detected in the raw water.
Less effective disinfection due to elevated turbidity as a result of surface water runoff following heavy rainfall in the catchment	4	4	16	Very high	Operational staff frequently report elevated levels of turbidity in the treated water following heavy rainfall, concomitant with lower residual chlorine levels. Water quality testing supports this, demonstrating increased detection of *E. coli* during these periods.
Ingress of contaminated water into the water distribution network following unsanitary pipeline repair practices	4	3	12	High	Rise in customer complaints recorded relating to dirty water in the vicinity of pipeline repair jobs. Water quality monitoring supports this observation, with elevated turbidity levels and reduced residual chlorine levels frequently detected following pipeline repair, which can cause widespread public complaint and has the potential to cause illness in the community.

2.2.1 Strengths

The main strength of this risk assessment approach is that it is a simple matrix that can capture different types of risk. The assessment would typically cover a more comprehensive set of hazards and hazardous events compared with sanitary inspections, including existing and potential hazards and hazardous events. The risk scoring can be used to prioritize remedial actions and to set alert or action levels. Like the sanitary inspections, risk scoring can be harmonized by using databases, risk scoring tools and predefined risk scores that are available and, in some countries, provided as guidelines or protocols by the regulator. Assigning risk scores requires expert judgement that may be facilitated by the databases and tools. However, as for sanitary inspections, the use of risk matrices is not a desktop exercise but requires the WSP team to collect and evaluate information about hazards and hazardous events on site.

2.2.2 Limitations

WSP teams find that the risk matrix may lead to extensive discussions over the likelihood of hazardous events and the severity of hazards. Sometimes it is difficult to apply a consistent method over the different hazards and hazardous events throughout the water supply chain. In particular, theoretical risks of chronic exposure to chemical hazards tend to be overrated in comparison with established acute risks of microbial hazards. Developing and agreeing upon well-defined definitions in advance of the risk assessment should aid in consistency. Evaluating the efficacy of control measures is not always easy, particularly controls in the catchment of source water for piped or non-piped distribution systems. The team's experience is usually dominated by nominal performance of control measures, unless there is a recent or documented history of hazardous events. Assessment of the frequency of infrequent hazardous events is therefore difficult. To obtain a robust risk assessment, it is important that the risk assessment team is knowledgeable about the water system and about hazards and hazardous events and incorporates the knowledge obtained at similar systems (i.e. via the checklists, databases, tools). Careful consideration of water quality data adds to the robustness of the risk assessment (see Box 2.3).

Box 2.3 Water quality monitoring to support the risk matrix

Water quality monitoring may inform the risk assessment in the WSP. A surface water supply using ozonation as the main disinfection process monitored water flow, ozone dose, residual and water temperature. Based on these parameters, the expected log removal of bacteria such as *E. coli* was very high (>5 log). Within the WSP risk matrix, the efficacy of this control measure was judged to be high. However, re-evaluation of the weekly monitoring data of the influent and effluent of the ozone reactor showed occasional breakthrough of *E. coli*, indicating very low log removal. Careful inspection of the ozone reactor showed a short circuit in the hydraulics of the reactor, which meant that a fraction of the water received very little ozone. In response, baffles were installed to improve the hydraulics. The water quality data provided the evidence base for the validation of the performance of this control measure and allowed for a more reliable risk assessment.

2.3 QMRA

QMRA offers a systematic way to use scientific information to help support water safety management decisions on a utility or regulatory level and prioritize remedial actions or research efforts (NRC, 2009). The numerical output of QMRA addresses the risk management questions in finer detail and allows for more precise comparison between risk management options compared with the qualitative or semiquantitative approaches introduced above. QMRA is a framework or mechanism that allows for quantitative scientific data to be interpreted in the context of estimated health outcomes in order to support water safety management (see Fig. 2.3).

QMRA is central in food safety management; the Codex Alimentarius Commission, established by the Food and Agriculture Organization of the United Nations (FAO) and WHO, has used it as the basis for international food safety standards and guidelines (Codex Alimentarius Commission, 1999). WHO has embedded QMRA in its water-related guidelines. For water safety management, the use of QMRA was first proposed in the early 1990s (Regli et al., 1991). Since then, QMRA has been used in different levels of detail for drinking-water, recreational water (outdoor and in pools), reuse of domestic wastewater or excreta in agriculture and many other forms of human water use. More detail will come with the expense of more time, more data and expertise; however, the outcome will result in more detailed understanding of the risks and control measures.

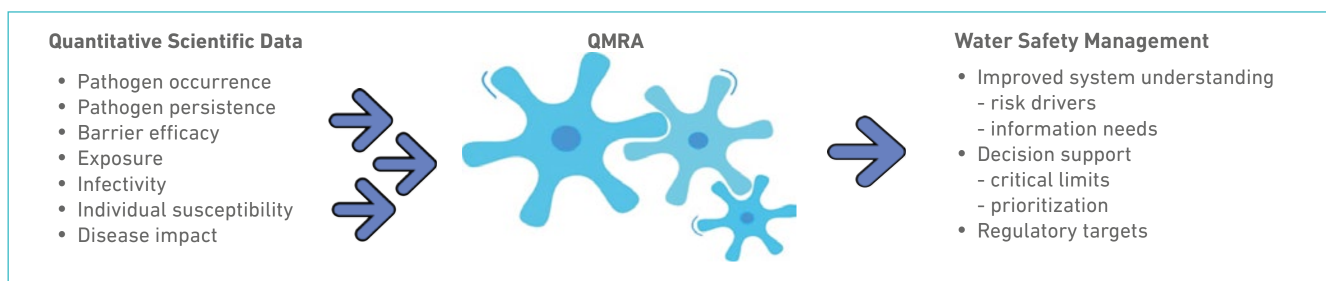


Fig. 2.3 QMRA: a tool for combining quantitative scientific data related to water-related disease pathways to support water safety management

QMRA is a formal four-step risk assessment process (Table 2.4), where each component of the assessment is explicitly quantified. The process of QMRA is described in the following sections, and the framework is described in detail in Chapter 3.

Table 2.4 Summary of the four-step framework for water-related QMRA

Step	Description
Problem formulation	The overall context (reference pathogens, exposure pathways, hazardous events and health outcomes of interest) of the risk assessment is defined and constrained in order to successfully target the specific risk management question that must be addressed.
Exposure assessment	The magnitude and frequency of exposure to each reference pathogen via the identified exposure pathway(s) and hazardous events are quantified.
Health effects assessment	Dose–response relationships (linking exposure dose to probability of infection or illness) and probability of morbidity and mortality (depending on the health end-point of the assessment) are identified for each reference pathogen.
Risk characterization	The information on exposure and the health effects assessment are combined to generate a quantitative measure of risk.

2.3.1 Problem formulation

In this step, the scope and purpose of the risk assessment are defined. This requires an interdisciplinary team to reach a balanced, robust and defendable risk assessment outcome that can be used to make decisions. Questions to be addressed by the team are: What risk management decision is required? What level of detail and what evidence base are adequate to support this decision? Which hazards (pathogens) and health outcomes should be considered? And which exposure pathways and hazardous events should be included? Depending on the scope of the risk assessment, different levels or tiers of QMRA, from screening to in-depth, can be applied. There is no sharp division between these tiers: they do not differ in nature, only in the level of data quantification and mathematical sophistication and the time, expertise and data required. The need of the risk manager in terms of level of detail and evidence base to inform risk management should determine the required level of sophistication. QMRA is well suited for a tiered approach, starting with a screening-level QMRA to determine the urgency of the perceived problem, to prioritize the risk of different water supply sites or scenarios and to determine the need for a more detailed study for a particular situation. This allows the effective allocation of resources to the sites or situations that give rise to the highest risk.

2.3.2 Exposure assessment

The aim of this step is to determine the frequency and magnitude of exposure to pathogens via the pathways and hazardous events defined during the problem formulation. On the one hand, quantitative information is needed – the pathogen concentrations in water sources used and fate of pathogens in barriers (e.g. water treatment processes) during normal and incident situations. On the other hand, data are needed on the exposure of humans – for example, 1) the size of the exposed population, 2) the nature of the exposed population (e.g. vulnerable groups) and 3) how often they are exposed (e.g. daily).

2.3.3 Health effects assessment

In this step, the health impact data for the identified hazards and the specific study population are compiled. This includes the type of health effects (including sequelae, or the severe, secondary and/or chronic health effects that occur following initial infection), the severity and duration of a disease or illness that may occur after ingestion of the pathogen and available information on the relationship between ingested dose and the probability that health effects (infection, illness, sequelae) occur (dose–response relationship). Also, the fraction and vulnerability of the population exposed may need to be considered.

2.3.4 Risk characterization

This step combines the information from the previous steps into an assessment of the probability of occurrence and severity of adverse health effects in the exposed population. This should be framed to address the risk management question defined at the outset of the QMRA – for example, the relative risk of different risk management scenarios or whether the water system meets a health-based target. Although calculated risks can be compared against a health target, users of QMRA should keep in mind that QMRA does not calculate actual disease outcomes, but provides a probability that disease may occur through the water system.

The time scale in which the risk is expressed may differ, from single-exposure events to all exposures in a year. The risk may be quantified in different end-points, including the probability of infection, probability of illness, expected number of illness cases and measures for burden of disease, such as disability-adjusted life years (DALYs) (see section 7.2). The DALY is the metric used in WHO guidelines for overall community health burden (WHO, in preparation). DALYs have the advantage that they allow weighting to be given to illnesses that lead to more serious health outcomes and comparison of different types of health risks.

It is important to include the variability (natural dispersion in a system, such as pathogen concentrations in a river) and uncertainty (lack of understanding and/or inability to measure) in all steps of the risk characterization. Sources of uncertainty in QMRA include extrapolation from dose–response data (although, unlike with toxic chemicals, many dose–response data are from human exposure), limitations of pathogen detection methods and estimates of exposure. To determine how the variability and uncertainty in the information at individual steps of the risk assessment affect the overall risk estimate, sensitivity analysis can be used.

The risk characterization is either deterministic (meaning that single values such as means are used to describe the variables used in the QMRA model) or probabilistic (meaning that statistical distributions are used to describe variables used in the QMRA model). In a deterministic QMRA, estimates of each of the QMRA model variables in the exposure and effects assessment are selected and combined to compute the resulting health risk. In a probabilistic QMRA, statistical distributions are used to describe the model variables, which reflects the stochastic (variable/uncertain) nature of most of the model variables more appropriately. The type of distribution selected considers a combination of knowledge of (pathogens in) water systems and of statistics. The health risk is computed by combining the statistical distributions, using Monte Carlo methods (Haas, Rose & Gerba, 2014). There is a wide range of software tools available to support these calculations. Dedicated QMRA models and software tools are increasingly available to aid this risk characterization step (Bassett et al., 2012). The need to conduct a deterministic, screening-level QMRA or a probabilistic, in-depth QMRA is primarily determined by what is needed to determine the best risk mitigation options. Chapter 9 provides more examples to illustrate how QMRA can support water safety planning and serve risk management and to illustrate the process of QMRA at different levels of detail.

An example of a (deterministic) QMRA to determine the safety of a surface water supply is given in Box 2.4.

Box 2.4 Example: QMRA to determine treatment performance targets following the WHO GDWQ

Problem formulation

A drinking-water supply is producing drinking-water from a river, using conventional treatment (coagulation/filtration) and ultraviolet (UV) disinfection. This river flows through a number of large urban centres and agricultural lands. Looking at the WHO GDWQ, the utility would like to demonstrate that the drinking-water meets the health outcome target of 10^{-6} DALY per person per year (pppy). Given the human and animal faecal sources in the catchment, the utility wants to include enteric bacteria, viruses and protozoa in the QMRA. It was decided to focus on *Campylobacter*, rotavirus and *Cryptosporidium* as reference pathogens for demonstrating that the target is being met. It was also decided to translate the 10^{-6} DALY into treatment performance targets, following the approach in the WHO GDWQ (Table 7.4 in the WHO GDWQ).

Exposure assessment

The utility has no data on the selected reference pathogens *Campylobacter*, rotavirus or *Cryptosporidium* in source water or on pathogen removal by the water treatment processes. Therefore, the utility made an effort to collate data on the occurrence of these reference pathogens in water systems with similar source water (sewage-impacted rivers with limited catchment protection) from the scientific literature. As there was uncertainty about how to translate these literature data to the specific site, the 95th percentile of the literature data from these source waters was selected as input for the QMRA. Limited data on rotavirus were found, so the utility used culturable enterovirus data instead. Similarly, data on the removal of the reference pathogens by coagulation/filtration processes and by UV were taken from scientific literature (review by Hijnen & Medema, 2010). For coagulation/filtration, the mean removal deduced from all literature data was used; for the UV disinfection, the local UV fluence setpoint (400 J/m^2) was used to translate the literature data to the local system.

For the consumption of (cold) tap water, no local data were available, so 1 L per person per day was used (as assumed in Table 7.4 of the WHO GDWQ).

	Data source	*Campylobacter*	Rotavirus[a]	*Cryptosporidium*
Source water (number/L)	95% of data from literature on sewage-impacted rivers	240	0.5	11
Coagulation/filtration (log removal)	Hijnen & Medema (2010)	2.1	3.0	3.2
UV (log removal)	Hijnen & Medema (2010)	>5	4.0	>3

[a] Virus concentration in source water is based on literature data on culturable enterovirus.

Health effects assessment

The dose–response relationships for *Campylobacter*, rotavirus and *Cryptosporidium* that are presented in the WHO GDWQ were used in this QMRA. In the absence of local data on disease burden (DALYs per case) and susceptible fraction of the population for these pathogens, these values were also taken from the WHO GDWQ.

	Data source	*Campylobacter*	Rotavirus	*Cryptosporidium*
Dose–response (probability of infection per organism)	GDWQ Table 7.4	0.019	0.59	0.2
Risk of illness given infection	GDWQ Table 7.4	0.3	0.5	0.7
Disease burden (DALY/case)	GDWQ Table 7.4	4.6×10^{-3}	1.4×10^{-2}	1.5×10^{-3}
Susceptible population (%)	GDWQ Table 7.4	100	6	100

Box 2.4 Example: QMRA to determine treatment performance targets following the WHO GDWQ (continued)

Risk characterization

The risk management question was to evaluate whether the water supply system was capable of meeting the health outcome target of 10^{-6} DALY pppy. A deterministic QMRA was conducted, based on the data collected from the literature. The main uncertainty is the validity of the use of the literature data for this specific water system. This led to the selection of the 95th percentiles of the concentrations of the reference pathogens as input values for the risk characterization. The treatment performance target was computed from the difference between the pathogen concentration in source water and the water quality corresponding to the health outcome target of 10^{-6} DALY pppy (see also above). This was compared with the estimated treatment performance (extracted from literature data; Hijnen & Medema, 2010), showing that multiple barriers in the treatment were needed and together were capable of producing water that meets the 10^{-6} DALY pppy target.

	Data source	*Campylobacter*	Rotavirus[a]	*Cryptosporidium*
Source water (number/L)	95% of data from literature	240	0.5	11
Pathogen concentration (number/L) in drinking-water corresponding to 10^{-6} DALY pppy[b]	–	1.1×10^{-4}	1.1×10^{-5}	1.3×10^{-5}
Required log removal to meet health-based target (performance target)	–	6.3	4.7	5.9
Calculated log removal	–	>7.1	7.0	>6.2
Coagulation/filtration	Hijnen & Medema (2010)	2.1	3.0	3.2
UV	Hijnen & Medema (2010)	>5	4.0	>3

[a] Virus concentration in source water is based on literature data on culturable enterovirus.

[b] From Table 7.4 of the WHO GDWQ.

2.3.5 Strengths

The main strength of QMRA is that it is an evidence-based, objective and transparent approach to provide, in greater detail, the quantification and discernment between risks compared with the other risk assessment approaches presented. QMRA supports system understanding: it considers all components in the water (supply) system from source to water user, providing valuable information on the effects of each component on the risk of human disease associated with exposure to waterborne pathogens (Medema et al., 2006; Ashbolt et al., 2010; Bichai & Smeets, 2012; Payment, 2013; De Keuckelaere et al., 2015). Owing to the inherent uncertainties associated with quantifying exposure, the greatest value of QMRA may not be in the final quantification of risk, but rather in the exploration of system variables and risk drivers to support water safety management. The results provide a scientific basis for evaluation of risk management priorities or control strategies. As such, QMRA can provide justification for allocation of resources to risk control measures (Bichai & Smeets, 2012). Although other risk assessment approaches can also provide justification for investments in improvements, QMRA can provide a more precise justification, which may be particularly useful when significant investments are required. Further QMRA enables the development of performance and specific technology targets, to determine whether or not microbial health outcome targets can be met. QMRA has made it possible to evaluate the health risks associated with waterborne pathogens in a broader context, such as risk trade-offs between pathogens and disinfection by-products (Havelaar et al., 2000) or arsenic (Howard et al., 2006), and waterborne disease burden estimates from QMRA are used in cost-of-illness estimates and cost–benefit evaluation of risk control strategies (see Case-study 4 in Annex A).

QMRA can be targeted towards specific risk management questions and be made “fit-for-purpose”, from screening-level using point estimates to in-depth using probabilistic descriptions for the model inputs.

2.3.6 Limitations

QMRA requires more technical knowledge and resources compared with the risk assessment approaches presented previously. As noted above, the level of quantification, mathematical sophistication, time, expertise and data required depend on the level of sophistication of the QMRA. Conducting a comprehensive probabilistic QMRA requires the most resources in terms of data and specialized personnel. In practice, such comprehensive QMRAs are conducted mainly in larger utilities in high-income countries where the public health agencies have promoted the use of QMRA by providing guidance or even direct support (e.g. Anonymous, 2001, 2005; USDA/USEPA, 2012).

A key limitation of QMRA is the limited availability of data on pathogen occurrence, fate and transport and removal by water treatment processes. Pathogen detection methods are becoming more widely available, but still require a high level of expertise to produce data of a reliable quality. When data are absent, QMRA has to work with assumptions. The use of default assumptions is well recognized in risk assessment (NRC, 2009). Where data are absent and assumptions are used, the tendency is to use conservative or worst-case estimates, to "be on the safe side". But worst-case estimates, by nature, may overestimate the risk, and it is not clear to the risk manager what the uncertainty of the calculated risk is – only that the uncertainty will be towards the lower-risk values (the nature of a worst-case assumption). To help the risk manager in making well-informed decisions, it is useful to provide a range of risks (interval estimate) that denote the variability and uncertainty in the risk estimate. This allows the risk manager to understand the nature and level of uncertainty in the analysis and the factors that drive this uncertainty. This helps the risk manager to decide whether the uncertainty is significant, or which uncertainties are significant, for decision-making. In the case of the screening-level QMRA, this can be achieved by using an average, worst and best case, to illustrate the range of the risks that can be deduced from the available information and the level of certainty that is embedded in the QMRA. An example of such a QMRA can be found in Medema et al. (2009). Alternatively, sensitivity analysis can determine the impact of variability and uncertainty in the variables on the certainty of the overall risk estimate. The nominal range sensitivity analysis method (Frey & Patil, 2002) is relatively robust; here, the value of each input variable in the QMRA model is varied, one at a time, along the confidence range of that parameter (e.g. average and maximum concentrations of a pathogen in water) to determine the effect on the final risk estimate. This procedure generates the range of possible values of the final risk estimate and indicates which of the variables contribute most to the uncertainty of the final risk estimate.

2.4 Selecting the appropriate risk assessment approach

The risk assessment approaches described in this chapter represent a continuum from simple to detailed and from expert judgement to evidence-based assessment of the risks (see Fig. 2.1). Table 2.5 summarizes the strengths, weaknesses, resources and expertise requirements for each of these risk assessment approaches.

A number of factors should be considered when selecting the risk assessment approach. One key consideration is to select the approach that provides information that is adequate for the risk manager as the basis for informed, evidence-based risk management decisions or to design risk management policies. Another key factor is to select the approach that fits in the context of human resources (personnel, skills, access to support institutions). Finally, the type of information (knowledge of the water supply system, types of hazards and hazardous events, exposure routes, water quality data on indicator organisms or pathogens, etc.) that is or can reasonably be expected to be available should be considered.

This does not imply that QMRA is appropriate only in settings with microbial risk assessment experts and systems with pathogen data. It is not the data that (should) drive the level of detail in the risk assessment; it is the level of detail required by the risk manager to make evidence-based, proportionate decisions on risk management measures or strategies. Risk management usually requires delicate trade-offs about the proportionate allocation of resources or about the level of protection of water safety versus other societal demands – as in the case of designating land use for riparian buffer zones to protect streams from agricultural runoff versus designating this land for agricultural use; or in the case of energy input for UV disinfection systems versus energy-saving; or in the case of chemical disinfection for pathogen control versus disinfection by-product formation.

Table 2.5 Comparison of risk assessment approaches to support water safety management

Approach	Strengths	Limitations	Resources	Expertise
Sanitary inspection (with water quality surveillance)	Simple, requiring the least amount of time and resources Enables easy identification of ongoing problems with the water supply and required improvement needs Can easily complement sanitary inspection with water quality data, and vice versa When applied broadly, can provide a regional or national snapshot of status of water supplies	Only visible hazards Each sanitary inspection is for a single time point Simplified hazard evaluation (limited to key risk factors) For an individual water supply system, does not discern priority level for unaddressed risk factors	Site visit Relevant sanitary inspection form/ checklist (Water quality assays: field or laboratory)	Qualified/trained inspector (Water quality assessment)
Risk matrices	Approach allows for a more comprehensive consideration of hazards/hazardous events than sanitary inspections Simple prioritization structure that allows different scenarios to be compared and supports identification and management of the most important risks	Limited precision for comparing hazards or hazardous events Based on expert judgement Sometimes may be difficult to agree on risk scores; when clear and robust definitions for the likelihood and severity categories are not developed and applied, there can be inconsistent risk scoring and imbalance between acute and chronic health effects	All of the above plus: Information for evaluating the frequency of various hazardous events and severity of hazards (e.g. hazard(ous event) databases, historical records)	Interdisciplinary WSP (or SSP) team with broad qualifications covering the supply chain from source to exposure (e.g. engineers, water quality experts, catchment experts)
QMRA	Quantitative outcomes for quantitative problems Direct input from statistical inference of observational data	Most complex, requiring the most expertise and data Data are limited for quantifying model inputs Uncertainty is often difficult to quantify and incorporate in risk outcomes Validity of default assumptions may be difficult or impossible to establish for study site	All of the above plus: Quantitative data or assumptions regarding pathogen occurrence, exposure and health impacts For more in-depth QMRAs, computational tools may be required	Risk assessor(s) relying on the expertise of the WSP team and all available data Risk assessors need to be knowledgeable about interpretation of microbial data sets and, for in-depth QMRA, modelling approaches, statistics, etc.

Higher trade-offs require more detail and evidence base in the risk assessment. An example of a risk assessment that required a high level of detail and evidence base can be found in the cost–benefit assessment of a new drinking-water regulation in the USA in Case-study 4 (see Annex A). Proportionate allocation of resources is also the basis of risk management questions, such as: which of the more than 1000 water supplies in country X are (most) at risk of faecal contamination, and thus where should priorities be placed for allocation of resources for risk reduction? Answering such questions can be done through sanitary inspections when conducted as part of surveillance as well as by QMRA (see the example from France in Case-study 3 in Annex A, with simple data from a water supply system survey and from scientific literature), depending on the evidence base required by the managing body and the availability of resources. More information about the interaction between risk manager and risk assessor to select, conduct and use the appropriate risk assessment is given in Chapter 9, and the case-studies in Annex A provide examples of how this has been applied in specific cases.

3 FRAMEWORK FOR WATER-RELATED QMRA

In this chapter, the four-step harmonized framework for undertaking QMRA in the water-related context is presented.

A range of different frameworks have been presented and applied for undertaking QMRA in water-related exposure settings. The earliest studies relied primarily on the framework developed for chemical risk assessments (NRC, 1983); subsequent work was undertaken to develop a framework that was specifically suitable for microbial risks (Soller, Eisenberg & Olivieri, 1999; Teunis & Havelaar, 1999; ILSI, 2000; USDA/USEPA, 2012). There are some differences in terminology, order and detail of content; however, the essential activities and components are equivalent. Different frameworks have been presented in each of the water-related guidelines (see Table 7.3 in the GDWQ [WHO, in preparation]; Table 2.2 in the GWEG [WHO, 2006a,b]; and Table 4.4 in the GREC [WHO, 2003] for comparison). Therefore, a single framework is now presented to harmonize across these guidelines. The four essential activities, as presented in Chapter 2, are referred to here as problem formulation, exposure assessment, health effects assessment and risk characterization (Fig. 3.1).

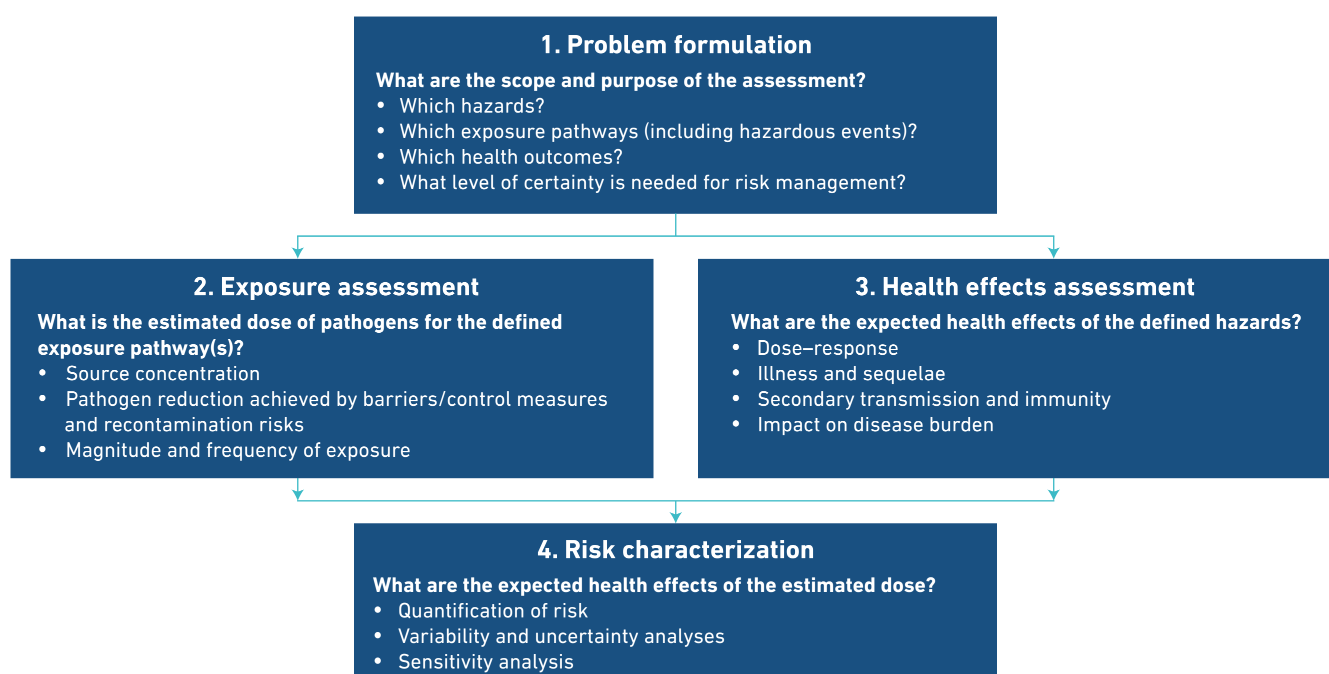


Fig. 3.1 Harmonized framework for water-related QMRA

3.1 Problem formulation

The first step in any QMRA is to define the purpose and scope of the investigation. The purpose will depend on the broader risk management context, the specific question that the QMRA is intended to address and the required level of certainty for appropriate risk management (see Chapter 5). The scope of the QMRA is defined by the identified hazards, exposure pathways and health outcomes to be considered:

- **Hazard identification:** It is not possible to consider all water-related human pathogens in a QMRA; therefore, *reference pathogens* are chosen that, if controlled, would ideally ensure control of all pathogens of concern. Reference pathogens should be selected taking into account local conditions, including relevance to the exposure pathway(s), source water characteristics and the incidence and severity of waterborne disease.
- **Identify exposure pathways:** The overall pathway from pathogen occurrence to human exposure is identified with a view to scoping the risk assessment and determining what will be included and excluded. Specific hazardous events (i.e. an incident or situation that can lead to the presence of the hazard) or scenarios that are to be included within the assessment in order to meet risk management objectives are identified.
- **Identify health outcomes:** The human health outcomes that are of interest are identified. Depending on the purpose of the assessment, the human health outcomes may include infection, illness, illness and sequelae, or a measure of disease burden that aggregates the impact of all of these outcomes (e.g. DALYs; see section 7.2).

3.2 Exposure assessment

The objective of the exposure assessment is to estimate the magnitude and frequency of exposure to pathogens via the identified exposure pathways and during hazardous events defined in the problem formulation.

Exposure assessment involves the following steps:

- **Define the exposure pathways** identified during problem formulation in detail, including points of quantification of pathogen sources; reduction (or recontamination) across natural, engineered and regulatory control measures/barriers; and mechanisms of exposure, generally as a set of exposure scenarios. This approach (described in more detail in section 6.1) facilitates harmonization of data interpretation and statistical methods across drinking-water, recreational water and wastewater reuse pathways.
- **Quantify each component of the exposure pathway** using the best available scientific evidence and an understanding of the expected variability and uncertainty associated with each model variable. This is addressed in detail in section 6.2 and Annex C.
- **Characterize the exposure** by quantifying the magnitude and frequency of exposure for the range of scenarios to be considered in the QMRA.

3.3 Health effects assessment

In the health effects assessment, the health impact data for the identified hazards and the specific study population are compiled. Special consideration may need to be given to vulnerable portions of the population (children, pregnant mothers, elderly and otherwise immunocompromised individuals) and to the fraction of exposed people in the total population. Depending on the health outcomes required for the assessment (identified during the problem formulation), the following components may need to be considered:

- **Dose–response:** Application of a dose–response model is the critical link between pathogen exposure and estimated health outcomes (either infection or illness). A model needs to be selected from the published literature that will be appropriate for the particular study. A review of relevant factors is included in section 6.1 and Annex D.
- **Probability of illness:** Not all infected individuals will develop symptoms. When using an infection-based dose–response model, it may be necessary to estimate the probability of illness once infected.
- **Probability of sequelae:** Sequelae are severe, secondary and/or chronic health effects that may occur following initial infection. Quantifying total disease burden involves giving consideration to the likelihood and consequences of these more severe health outcomes for the given population.
- **Disease burden:** The DALY is the metric used in WHO guidelines for overall community health burden. For water-associated diseases, it incorporates the total impact of all health outcomes listed above on the exposed population. The advantage of using DALYs is that it allows the consideration of different impacts on both the quantity and quality of life. It has been adopted for both the drinking-water and wastewater guidelines as the measure of setting health-based treatment targets. To learn more about health-based targets including DALYs,

the reader should refer to the WHO guidelines for drinking-water (WHO, in preparation), recreational water (WHO, 2003) and wastewater (WHO, 2006a,b), the WHO estimates of the global burden of foodborne disease (WHO, 2015b) and section 7.2.

- **Secondary transmission and immunity:** Secondary transmission (additional cases in the community due to person-to-person contact) and immunity (inability of an individual to become infected due to protection via his or her immune system) are important for the overall evaluation of the health impacts. These factors can be assessed within QMRA using what are called dynamic risk models.

3.4 Risk characterization

The exposure and health effects assessments are combined and calculations undertaken in order to quantify and characterize the risk. The content of the risk characterization is driven by:

- **The purpose of the risk assessment as defined during problem formulation:** The purpose will drive what model conditions are simulated and run for the calculations. In a very simplistic model, this may require only a linear quantification of risk. In a more complex model, specific scenarios and conditions need to be selected for the calculations in order to meet the purpose of the assessment.
- **Quantitative measures of risk:** The risk may be quantified in many different metrics from the information collected during the exposure and health effects assessments, including the probability of infection, probability of illness, expected number of illness cases and DALYs. The time scale of risk may be for a single exposure, a series of independent exposures or a year. The population may be the total population or the exposed fraction of the population.
- **Variability and uncertainty (see Chapter 4):** Depending on the scope of the risk assessment, the risk may be characterized by a single point estimate, such as the mean, minimum or worst case, or by a probability distribution (distribution that takes into consideration the range of likely values and the probability of each of those values occurring). When uncertainty (lack of knowledge) and variability (how a value or attribute varies) have been explicitly quantified for the model inputs, the risk characterization will involve quantifying the distribution of risk. See Chapter 4 for a description of variability and uncertainty concepts and Annex B for an example demonstrating the incorporation of variability and uncertainty within QMRA.
- **Sensitivity analysis (see Chapter 8):** Sensitivity analysis (the investigation of how variability and uncertainty in the input parameters influence the variability and uncertainty in the risk output) can be used to explore how the model components or variables, such as pathogen concentrations, efficacy of intervention measures, dose–response parameters, morbidity ratio, etc., interact and which is most important to the outcome. In particular, sensitivity analysis allows the most important sources of variability and uncertainty to be identified, which can be used to target where management should focus control measures and additional data collection. A more detailed description of sensitivity analysis is given in Chapter 8.

4 | VARIABILITY AND UNCERTAINTY

Before presenting the individual steps of the QMRA in Chapters 5–8, this chapter describes how variability and uncertainty can be incorporated into QMRA.

Variability refers to how system elements change over time and space. Additional data collection does not reduce variability, but can better characterize how elements vary. Uncertainty refers to a lack of knowledge regarding system elements and, ultimately, risk. Uncertainty forces the risk manager to judge how probable it is that the risks are overestimated or underestimated and select an appropriate degree of conservatism.

Variability and uncertainty are common elements of all types of risk assessment. An understanding of the impact of variability and uncertainty in the input parameters on the risk output of the QMRA allows the evidence to be weighed when designing an appropriate risk management response to the health risk.

4.1 Variability

QMRA models describe natural systems, environmental processes and their interactions with respect to pathogens from source to exposed people. Each component of the exposure assessment and health effects assessment is likely to vary both spatially (between locations, individuals or pathogens) and temporally (over time), leading to variability in the risk to human health. The risk will therefore vary between sites (e.g. bathing may be considered safe in one context, but unsafe in another), over time (for any given exposure pathway and context, risks will fluctuate; over the long term – from seasonal fluctuations to climate change trends – and the short term, particularly in response to specific events such as rainfall or engineering events such as treatment failure/poorer performance) and between individuals (different subsets of the population may be exposed to different levels of risk and may show different levels of sensitivity to the pathogens). Examples of common sources of variability in the context of water safety management are:

- **Pathogen concentrations in source water:** Pathogen concentration in surface water and groundwater will vary between and within sites depending on many factors, including the catchment size, geohydrological conditions, topography, climate, land use and the associated faecal sources upstream of the sampling location, and the presence of (engineered and non-engineered) interventions to reduce discharge of pathogens to surface water and groundwater. For a single site, the concentration will vary over time, depending on the incidence of infection in faecal sources, excretion dynamics and quantity (volume/mass); environmental events (e.g. rainfall-induced runoff or infiltration) that mobilize pathogens; and environmental conditions (including temperature, sunlight and biotic [predator] activity).
- **Pathogen concentrations in treated sewage effluent:** Pathogen concentration in treated sewage effluent will vary between and within sites, depending on the size of the population contributing to the effluent, the dilution, the prevalence of infection in that population and the performance of the specific treatment barriers in

place. For a single site, the concentration will vary over time, depending on the incidence of infection (at times seasonal) and fluctuations in hydraulic loading and treatment performance.

- **Performance of engineered treatment barriers:** The performance of engineered treatment barriers will vary between sites and over time, depending on many factors, including inflow water quality, quantity and variability, chemical dosing, specific process design characteristics and operation/management decisions, and within sites, as a result of local fluctuations and process events.
- **Performance of natural treatment barriers:** The performance of natural treatment barriers will vary between sites, owing to the unique local environmental conditions, and within sites over time, as a result of seasonal changes and response to specific events such as rainfall or snowmelt.
- **Exposure volumes and frequencies:** Exposure volumes and frequencies will vary between and within individuals, depending on climate, age, lifestyle and cultural factors, and within sites over time, as a result of seasonal changes and participation in exposure activities.
- **Dose–response:** Infectivity will vary between and within individuals, depending on immune status, age and health factors, and between and within pathogens, owing to differences in strain virulence and exposure pathways, the physiological state of the pathogen and other pathogen-specific parameters (e.g. extent of microbe aggregation).

These sources of variability lead to variability in health risk. Waterborne outbreaks have occurred because of the peaks in pathogen concentrations in water systems (Risebro et al., 2007), with the Milwaukee outbreak of *Cryptosporidium* as a notorious example (MacKenzie et al., 1994). Variability therefore needs to be considered in the risk assessment. Section 4.3 explains how variability can be incorporated into QMRA.

4.2 Uncertainty

Uncertainty is inherent in all types of risk assessment (Morgan & Henrion, 1990). Quantifying each of the model inputs is a challenge, as knowledge about the system is generally imperfect, and sometimes knowledge about model inputs is limited or absent. The validity of the risk assessment relies on how well each selected value (or distribution) represents the true value (or distribution) of each of the model variables. In general, the more information that is available and the more data that are collected, the lower the uncertainty.

Building QMRA models involves subjective choices and assumptions to fill in the gaps of limited data sets. For this subjective selection of information and assumptions, expert opinion plays an important role in selecting the best available information and in interpreting potential risk (management) implications. For the risk manager, it is critical to be able to evaluate the uncertainty associated with these assumptions, even when they cannot be easily quantified. A precautionary approach to risk management generally drives the selection of conservative or worst-case model assumptions. It is important that the level of safety and protection required by risk managers be considered in the problem formulation stage and that the impact of such selections on the risk calculations be assessed during risk characterization. Transparent evaluation can ensure that assumptions are appropriate to the purpose of the QMRA and not unreasonably conservative.

Some common uncertainty considerations that arise in water-related QMRA studies include:

- **Uncertainty due to absence of specific information:** In many instances, data for one or more of the steps in the QMRA from the water system under study are missing or inadequate. It is quite common in QMRA to use general information (e.g. from national statistics or international literature on microbial data) or data from similar systems to fill this gap. Although this allows the risk assessor to perform studies in data-limited situations, it comes with potentially greater uncertainty about the validity of the information for the system under study and about how representative the data are for the specific study site. Given this uncertainty, conservative values are often selected. For example, the 95th percentiles of national and international pathogen concentrations in domestic wastewater are used as the basis for QMRA for wastewater reuse in the Australian guidelines for water recycling (NWQMS, 2006).
- **Uncertainty regarding the representativeness of experimental data:** The risk assessor may often have reason to question whether the local data are representative. For example, the analyst may have a small data set of pathogen counts from the water system of interest, which consists of a series of "non-detects". Owing to some prior experience or knowledge (e.g. the range of pathogen densities expected given catchment sources or data on the variability in water quality obtained from microbial indicators), the analyst may question whether the data set truly reflects the underlying pathogen concentrations. Alternatively, where microbial surrogates are used to describe barrier efficacy, the analyst must question how well the surrogate data are expected to represent the

pathogen concentration and behaviour for the specific system. The decision to rely on these data, or to rely on alternative data, must be justified, and any limitations explicitly defined.

Consideration of this uncertainty is relevant for understanding what the risk estimates represent. At the same time, it helps in the verification of the QMRA by evaluating the (strength of the) information against the available scientific knowledge. Ignoring all prior knowledge from literature studies, other data sets or epidemiological experience in favour of local data alone is irrational, particularly if the local data set was extremely small or from an unknown source. The data themselves are subject to many uncertainties, including influences from random sampling, method recovery and the fraction of detected pathogens that may still be human infectious (see Annex C). It is therefore desirable to be able to use small local data sets, but then also to consider and preferably test the importance of any perceived inadequacies.

- **Uncertainty regarding selection of statistical distribution:** Probabilistic QMRAs rely on statistical distributions to describe model variables; however, microbial data sets are rarely large enough to demonstrate the full range of variability. The basis for the selection of statistical distributions (such as the Poisson or lognormal distribution) in a probabilistic QMRA should be well understood and explicitly stated. A range of distributions may often provide a similar fit. Although the risk assessor may believe that the selected distribution is the most appropriate choice given the available data and understanding of the system, there may be equally valid alternative choices. In such circumstances, it would be reasonable to undertake the QMRA with each alternative distribution so that the influence on the outcome could be assessed. If distribution choice proves to be important to the risk outcome, further data collection to understand which distribution is likely to be more representative would be justified.

The sources of uncertainty for water-related QMRA can be classified into two categories: model assumption uncertainty and parametric uncertainty:

- **Model assumption uncertainty:** In constructing or selecting, adapting and applying an existing QMRA model, it is necessary to make simplifying assumptions about the study system. Simplifications may include the selection of parametric distributions to describe variability (e.g. the normal distribution defined by the parameters mean [μ] and standard deviation [σ]; or gamma distribution defined by scale [λ] and shape [ρ] parameters), barrier reduction models (e.g. where the reduction in microbial concentration across a barrier is defined by a single $\log_{10}$ reduction or a distribution of $\log_{10}$ reductions; and where the total $\log_{10}$ reduction of a treatment train is assumed to be the sum of the $\log_{10}$ of the individual components, where a conservative $\log_{10}$ reduction credit for each barrier is selected; see section 6.2.2.3) and dose–response models (e.g. for single-hit models, where the probability of an individual microorganism passing all host defences and achieving infection is represented by a parameter *r*). These types of model simplifications and assumptions are necessary, but there is always some uncertainty as to how well these models represent the reality of the environmental system.
- **Parametric or statistical uncertainty:** Once a model is selected, the parameter values are estimated from, or fitted to, observational data. Three aspects of the data will influence the uncertainty in the estimated parameter values:
  1) **Sample size:** The number of observations. In general, as the size of the data set increases, the uncertainty in the parameter values will decrease.
  2) **Measurement uncertainty:** Variables that are modelled in QMRA are often difficult to measure and may be estimated or inferred rather than directly measured (e.g. interpretation of pathogen data discussed in sections C1 and C2 in Annex C).
  3) **Data representativeness:** The data set will describe only the specific conditions at the time the observations were made. It is important to consider how well the data set represents the conditions of the risk assessment – for example, considering whether the data set has samples collected under similar conditions (e.g. flow, season and climate); whether the sample was random, event driven or directed; and whether the sample is representative of the likely exposure pathway under consideration.

Table 4.1 includes a summary of these categories of uncertainty and some examples of the type of factors that would be identified through a structured consideration of each category.

Table 4.1 Examples of categorization of uncertainty associated with QMRA model input[a]

Examples of model inputs	Model assumption uncertainty	Parameter uncertainty		
		Sample size	Measurement uncertainty	Data representativeness
	How sure is the analyst that the model selected is appropriate?	What is the uncertainty in the predicted parameter values given the data?	How well do the observations describe the constructed model and target variable?	Do the data represent a random sample? Were the data collected under a specific set of conditions? How representative are those conditions of the risk assessment?
Concentration of noroviruses in surface waters assumed to be gamma[b] distributed (described by parameters ρ, λ)	Is gamma an appropriate choice? What if the concentration was bimodal (e.g. representing seasonal changes, or including a subpopulation clumped with particulates)?	Data set consists of 14 surface water samples. What is the uncertainty associated with predicting ρ, λ when the gamma distribution is fitted to the data set?	Analysis was undertaken by PCR and MPN estimation. What was the uncertainty associated with the MPN estimate? Was method recovery accounted for? How representative was the PCR measurement of infectious noroviruses?	Samples were collected every 14 days during winter. Does this represent the full range of variability for the water source? Given the assumption that norovirus concentrations in surface water are highest in winter (based on scientific literature), does the selection of the winter period for sampling introduce sufficient safety in the QMRA?
Inactivation of viruses[c] on food crops assumed to follow first-order inactivation kinetics $C_t = C_o e^{-kt}$ with inactivation rate = k	Is first-order decay a reasonable assumption? What if decay was biphasic due to clumping or resistant subgroup?	Data consisted of 10 replicate food samples collected on days 0, 1, 5, 10 and 15 following irrigation. What is the uncertainty associated with k based on these data?	What was the recovery and consistency of the method? What was the influence of counts and dilutions on the MPN or density estimate?	Data were collected using bacteriophage; would they be more or less persistent than human enteric viruses? Experimental conditions were under high temperature and humidity; how representative were those conditions of the crop irrigation conditions?
Dose–response modelled using Beta-Poisson model fitted to clinical rotavirus data[d]	Is the Beta-Poisson the appropriate model choice? What would be the impact of an alternative model?	What is the uncertainty in the prediction of α and β from the clinical data? Are these parameter values relevant to children, the most vulnerable group?	What were the uncertainties in the predictions of dose and infection associated with the clinical data set?	How representative of the whole population were the subjects (healthy adults)? How representative were the rotavirus strains used in the clinical study for environmental strains? How representative is the rotavirus model for describing dose–response associated with other enteric viruses?

MPN: most probable number; PCR: polymerase chain reaction

[a] These are hypothetical examples to explain concepts in section 4.2.

[b] See Box C.2 in Annex C.

[c] See section C6.2 in Annex C.

[d] See Annex D.

4.3 Accounting for variability and uncertainty in QMRA

The selected method of accounting for variability and uncertainty in QMRA is a strong driver of the mathematical complexity of the study. The more factors that are explicitly and quantitatively accounted for, the more detailed the analysis. Variability and uncertainty can be accounted for within QMRA using point values (deterministic) or distributions (stochastic). As these different inputs are combined, they will have an impact on the way the risk is characterized and on what is represented by those risk estimates.

The overriding principle associated with the incorporation of variability and uncertainty in QMRA is to adopt a tiered approach. The tiered approach to system assessment is illustrated in Fig. 4.1. The level of detail in the analysis should be as simple as possible, while achieving the objective of the QMRA defined in problem formulation. The initial screening-level assessment will typically involve point estimates of each input variable at their most likely, best-case or worst-case value. The limits of the risk outcomes will then inform whether further analysis is needed. For example, if all input variables are set equal to their worst-case values and the risk is still quantified to be acceptable relative to the target, then it is reasonable to assume that the pathway is safe and there is no need for further analysis. If all input variables are at their expected values and the risk is deemed to be very high, then management action is clearly necessary. Alternatively, if the point value analysis yields a result close to the target, then a tier 2 analysis is justified to explore in more detail the variability in risk. This is illustrated with a detailed numerical example in Fig. 4.2. See section 9.2.2 for additional information.

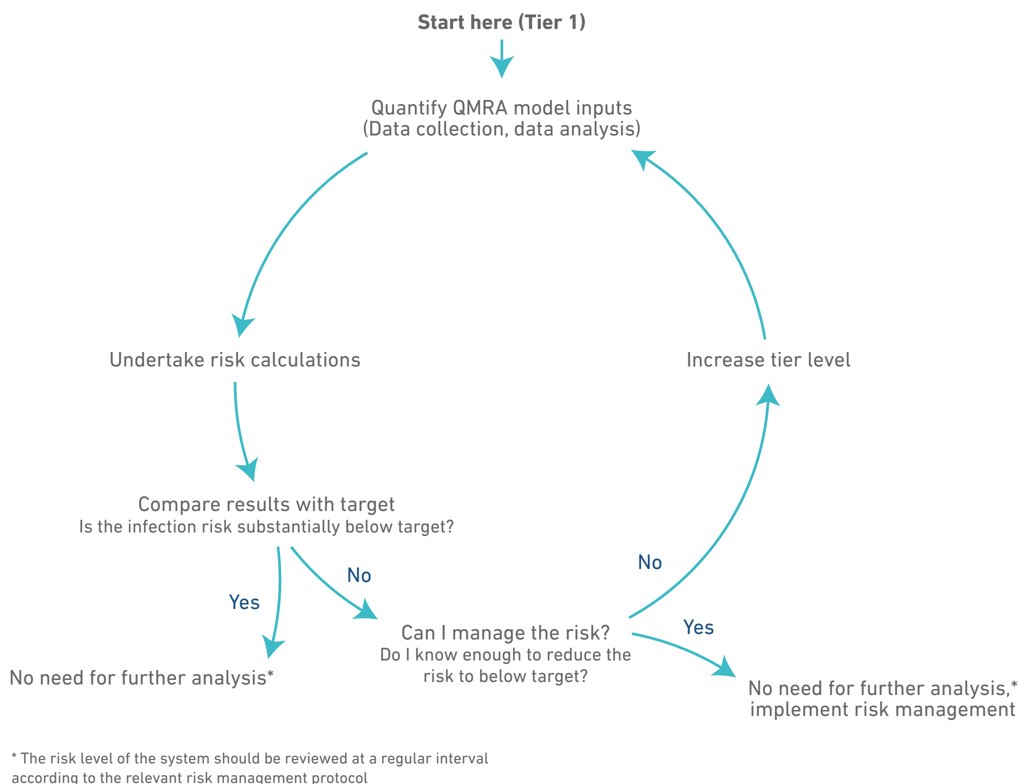


Fig. 4.1 Iterative approach to implementing QMRA for risk management; increase in tier relates to gathering more/ local data to reduce uncertainties (Medema et al., 2006)

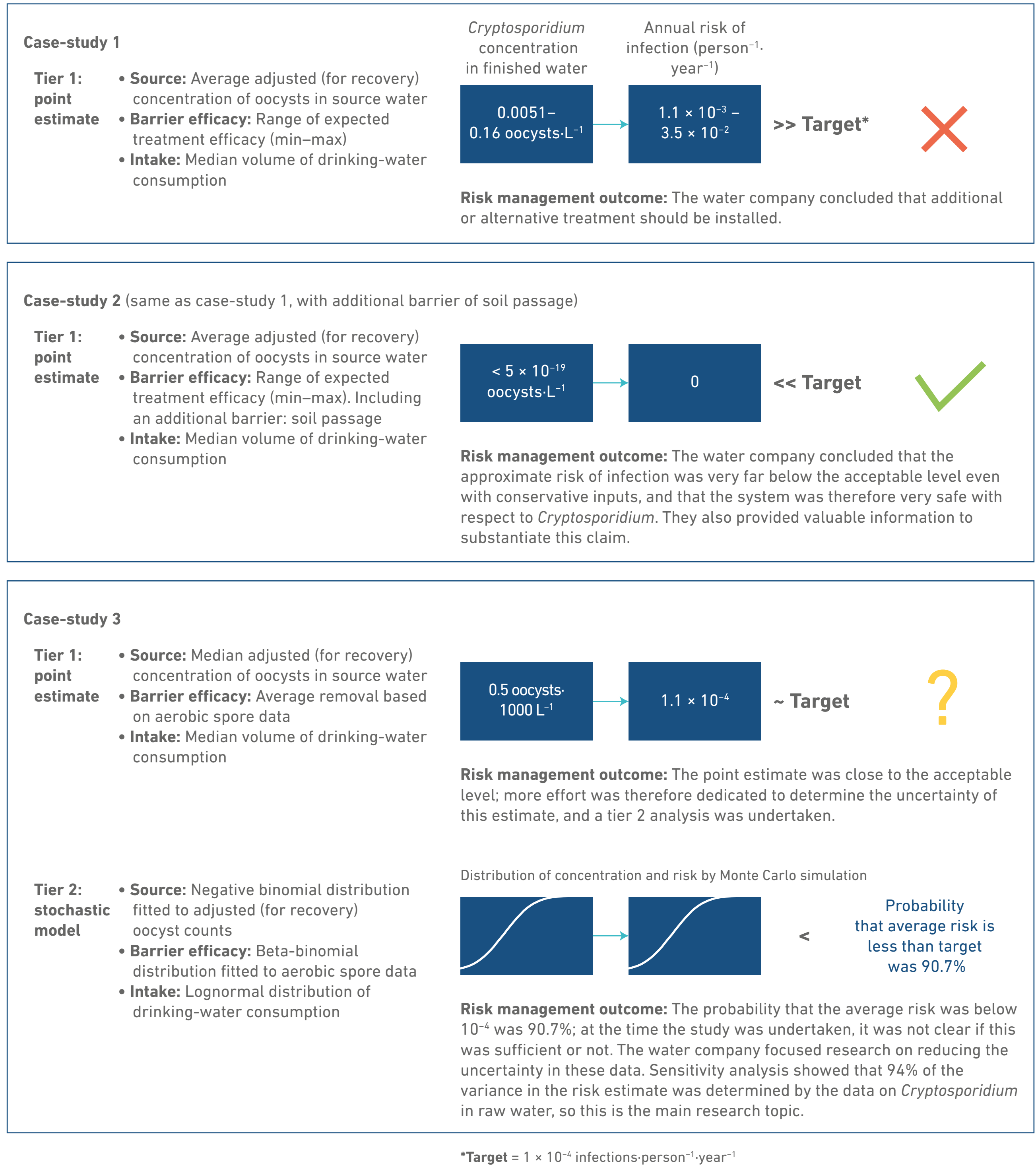


Fig. 4.2 Summary of tiered approach to QMRA for system assessment of drinking-water systems (based on data from Medema et al., 2003)

An illustrative example showing the stepwise incorporation of variability and uncertainty in QMRA is given in Annex B. There are many ways to account for variability within a risk assessment, but they are simplified here into two generalized approaches:

1) **Point estimates (deterministic):** Statistical descriptors of input variables, such as the arithmetic mean, 5th and 95th percentiles, and minimum and maximum observed values, can be used to explicitly account for variability (see Annex B, Part A). Scenario analysis can also be used, where the full range of variability in each of the model inputs is broken down to a set of defined conditions, each represented by a point value (Annex B, Part C). The risk can then be calculated for each set of conditions and used to provide an overall picture of the range of expected risks.
2) **Distribution (stochastic):** The full distribution of variability in the risk estimates can be quantified by describing each model input using a probability density function (PDF) (Annex B, Part D; Annex C, Box C.2). During risk characterization, the full probability distribution of risk is then quantified either theoretically or by simulation using Monte Carlo analysis (Annex C, Box C.2).

In the same way, uncertainty within a risk assessment can be described in many ways, which are simplified here into two generalized approaches:

1) **Point estimates (deterministic):** Uncertainty can be explored using point estimates for concentration of microbes and making assumptions about the available data. For example, what if the peak source water concentration was 10 times higher than what was measured? Or what if a different data analysis model were to be applied to the same system? Such exploration of the data and consideration of data uncertainty provide valuable insight to better understand the estimate of overall risk.
2) **Distribution (stochastic):** In addition to describing variability in model inputs, probability distributions can also be used to describe the uncertainty in model parameters (Annex B, Part E). The overall risk estimates and associated uncertainty can then be calculated using a Monte Carlo simulation (referred to as "second-order Monte Carlo", as parameters are randomly sampled first, followed by the model variable).

Although variability assessment and uncertainty assessment are undertaken during risk characterization, it is essential to note that they need to be considered for each of the model inputs. All scientific data and associated inferences carry different sources of uncertainty. Understanding the nature of these variabilities and uncertainties is important for the interpretation of calculated measures of risk. A particular emphasis is placed on these factors in Chapters 6 and 7.

In practice, there is great diversity in published studies regarding which sources of variability and uncertainty are considered and how they are accounted for in water-related QMRA. Even within a given study, the extent to which variability and uncertainty are accounted for is often not consistent throughout the model. For example, variability in source water concentration may have been considered, but not variability in treatment or each barrier in treatment. Also, the assessor may have considered parameter uncertainty for dose–response but not for source water pathogen distribution. Any calculations are a factor of the input assumptions, and the uncertainty and variability in the calculated risk (and any associated sensitivity analysis) are representative only of those factors that were defined in the inputs. Although it is not possible (or necessarily beneficial) to quantitatively account for all sources of variability and uncertainty in the QMRA, and sometimes it is not easy to distinguish them, it is important to consider those factors that are relevant to the objective of the study.

5 PROBLEM FORMULATION

In problem formulation, the overall context of the risk assessment is defined and constrained in order to successfully target the specific risk management question that must be addressed.

5.1 Interaction between risk manager and risk assessor

The first principle to keep in mind is that in order for the QMRA to be useful, it is critical that there is a transparent dialogue between the risk manager and the risk assessor throughout the assessment process. This begins with scoping and planning during the problem formulation stage to ensure that the risk assessment addresses the needs of the risk manager. In this interaction, it is important to acknowledge that there is a clear distinction between the role of the risk assessor and the role of the risk manager. The outcome of the risk assessment process should in no way be influenced by the preference of the risk manager. The important elements of the planning and scoping phase are presented in Box 5.1 as a series of questions to be pursued. Increased emphasis on planning and scoping has been shown to lead to risk assessments that are more useful and better accepted by decision-makers (USEPA, 2003, 2004). In the problem formulation phase, the scope is translated into a conceptual model and analysis plan. Technical limitations, such as limited availability of data and the need for assumptions and for process simplifications, will arise in this phase. It is important for the risk manager to be involved in this scoping phase to check the plan and perceived outcomes against the risk management needs. Also, during the process of the QMRA, as the information develops, the analysis plan may evolve. Therefore, communication between the risk assessor and the risk manager remains important to ensure that the utility for risk management "remains the major impetus" (NRC, 2009).

Box 5.1 Questions to help scope and plan a useful QMRA

a. What is the problem to be investigated, and what is its source?
b. What are the possible opportunities for managing risks associated with the problem? Has a full array of possible options been considered, including legislative requirements?
c. What types of risk assessments and other technical and cost assessments are necessary to evaluate existing conditions, and how do the various risk management options alter the conditions?
d. What impacts other than health threats will be considered?
e. How can the assessments be used to support decisions?
f. What is the required time frame for completion of the assessments?
g. What resources are needed to undertake the assessments?

Source: NRC (2009). Reprinted with permission from the National Academies Press, Copyright 2009, National Academy of Sciences.

5.2 Fit-for-purpose

An essential consideration at the initial stage is the required level of complexity needed to support the intended risk management outcome. Will a screening-level, deterministic risk assessment generate adequate information, or will a probabilistic risk assessment provide the information needed to support the risk management decision? It may be that an initial screening-level risk assessment is done to highlight a problem or eliminate a concern or that the screening-level risk assessment is a first crude test of the level of risk. The requirements of the risk manager may suggest the use of worst-case scenarios or more plausible upper bounds or representative central estimates, or perhaps even several of these, to gain an understanding of the effect on the uncertainty of the final risk estimate. If such a screening-level risk assessment indicates a low risk, even in the worst-case estimate, steps to further mitigate the risk are not justified (see Fig. 4.1 in Chapter 4). Often screening begins with worst-case estimates and then moves to increasingly more realistic point estimates as warranted. Conversely, if the risk estimate is unacceptably high, immediate remedial actions are appropriate, especially if the costs of these actions are low (Cullen & Small, 2004). However, in many cases, the costs of actions are substantial and the risk estimate is based on assumptions of which the validity is uncertain. This calls for more advanced information collection and risk assessment to support adequate decision-making. An in-depth QMRA requires further information collection and incorporation of the variability and uncertainty of the information into a comprehensive QMRA model. Using such a probabilistic approach allows a good understanding of the drivers of the risk and the most significant contributors to the overall variance in the risk estimate. This helps to weigh the feasibility and costs of collecting further information against the feasibility and costs of interventions. Table 5.1 gives the general characteristics of the different levels of QMRA. This may suggest a clear distinction between different levels of QMRA, but in practice it is a continuum, as represented by Fig. 4.1.

Table 5.1 Levels of QMRA

Level of QMRA	Characteristics
Screening	• Provides a broad overview of level of risk or safety: highlights water safety concerns or eliminates insignificant concerns • Provides crude understanding of the drivers of the risk • Collection and analysis of readily accessible, existing information about the water system, contamination sources, etc., such as in sanitary surveys • Rapid • Low cost • Tendency to use worst-case estimates on pathogen occurrence and barrier efficacy • Highly dependent on assumptions
Advanced	• Provides more detailed information on possible health risks • Provides better understanding of the drivers of the health risk and relative importance of barriers • Additional collection of more specific data and information, such as data on microbial contamination and the variation therein • Intermediate time span • Intermediate cost • Tendency to use best (point) estimates • More objective and reliable due to use of more specific information
In-depth	• Provides comprehensive understanding/robust definition of health risks • Provides comprehensive understanding of the drivers of the health risk and relative importance of barriers • Additional collection of more specific data and information, such as data on pathogen occurrence, barrier efficacy and the variation therein • Longest time span • Highest cost • Tendency to use probabilistic estimates • Most objective and reliable due to use of more specific information and probabilistic approach to incorporate variability and uncertainty

Source: adapted from Abrahams et al. (2004)

The usefulness of pursuing a more sophisticated QMRA depends on several factors. Reasons for not conducting further QMRA modelling include the following (Cullen & Small, 2004):

- immediate action is needed in response to an immediate concern;
- the results of the screening-level QMRA indicated that the risk is insignificant;
- the cost of assessment is higher than the costs of remediation; and
- the estimates/assumptions are highly uncertain, and more sophisticated analyses could further compound errors and incorrect results.

In contrast, reasons for conducting further QMRA modelling include the following (Cullen & Small, 2004):

- to avoid costs of remedial actions or other consequences based on poor or biased estimates;
- to increase objectivity and reduce dependency on assumptions;
- to support more sensitive cost–benefit analysis;
- to allow more rigorous comparison of alternatives;
- to provide a systematic consideration of all sources of variation and uncertainty; and
- to conduct sensitivity analysis and determine the validity of the outputs.

5.3 Including risk management options in the QMRA

A useful approach to facilitate the use of QMRA to guide decision-making is to include several alternative risk management options in the risk assessment and evaluate their impact on the estimated risk. The risk manager could specify these different options in the scoping and planning phase of the risk assessment. These could be alternatives within a specific risk management option, such as different levels of dosing (e.g. chlorine or UV fluence) in a water disinfection process. They could also be a comparison of different risk management scenarios, such as comparing the impact of wastewater treatment options with different irrigation systems and various withholding periods for irrigated crops. Given the current state of the art and knowledge, QMRA is particularly applicable for scenario evaluation and relative risk assessments. After evaluating the various options to reduce risk, the risk manager can select the preferred option. The value of this approach is demonstrated in Case-study 2 in Annex A (Weir et al., 2011). Pathogen risks can then be weighed alongside other important economic, social and cultural factors to support an overall cost–benefit analysis for selection of the best option.

5.4 Defining the scope of the QMRA

5.4.1 Hazard identification

Hazard identification is predominantly a qualitative process intended to identify microorganisms of concern (FAO/WHO, 2003). It is not possible to consider all human enteric pathogens in a QMRA; therefore, reference pathogens are chosen that are of particular relevance to the exposure pathways and context of the individual risk assessment. Reference pathogens are intended to provide a conservative model for the risk assessment: that is, if the reference pathogen is controlled, it is assumed that other important pathogens within each class would also be controlled. When considering human enteric pathogens, at least one bacterium, one virus and one protozoan are typically recommended in order to cover the range of behaviours in the main enteric pathogen groups. Inclusion of a representative helminth (e.g. *Ascaris*) is recommended for wastewater reuse and sanitation–based scenarios. Specific studies may require more detailed consideration of variability between pathogens within a particular group, in which case more than one reference organism should be selected (e.g. it may be relevant to model the highly chlorine-resistant oocysts from the parasitic protozoan *Cryptosporidium* alongside the less resistant but more numerous *Giardia* cysts in freshly contaminated waters). Both the GDWQ (WHO, in preparation) and the GWEG (WHO, 2006b) rely primarily on rotavirus, *Campylobacter* and *Cryptosporidium* as reference pathogens. However, the guidelines recommend that local considerations need to be accounted for, including the following:

- Epidemiological information on disease prevalence and outbreaks via the local water pathway(s): Which organisms are of most relevance for the population that is exposed? In many situations, the specific microbial agents responsible for disease may be unknown; however, it is reasonable (for the purpose of selecting appropriate reference pathogens) to infer the predominant etiological agent from disease symptoms.
- Scientific evidence on pathogen persistence and infectivity: which organisms are likely to be the biggest threat for the particular exposure pathway?

- Severity of disease outcomes: Which organisms are likely to be the most significant in contributing to overall disease burden?

Characteristics of reference pathogens commonly used in QMRA are summarized in Table 5.2. The availability of dose–response and disease burden data may also be a reason for the selection of the reference pathogens. In studies where no specific dose–response data were available for the selected reference pathogen, the conservative assumption has been to apply the dose–response data for the most infectious representative of the pathogen group (e.g. to use the dose–response data from rotavirus to describe the infectivity of enteroviruses, as is done in the Dutch Drinking Water Act in the Netherlands; see Box 9.1 in Chapter 9).

5.4.2 Exposure pathways

The overall exposure pathway is identified systematically from source to exposure, including pathogen sources, environmental fate and transport barriers, environmental or engineered barriers, and human activities that may lead to different levels of exposure. In settings where raw sewage and raw faecal waste are still widely prevalent, they should be addressed as pathogen sources. It is important to include any potential hazardous events or scenarios that are to be considered in the assessment within the overall exposure pathway in order to meet the identified risk management objectives. It may be helpful to construct a diagram identifying all system components that need to be quantified in the risk assessment.

5.4.3 Health outcomes

The human health outcomes that are of interest are identified. Depending on the purpose of the risk assessment, the human health outcomes may include infection, illness, illness and sequelae, or overall measures of disease burden that aggregate the impact of all of these outcomes (e.g. DALYs, see section 7.2). The health outcomes to be included in the analysis will depend on the purpose – for example, if the objective is to compare the calculated risk with a regulatory target, the requirements of the regulation (annual probability of infection or DALYs) will drive the selected health outcomes. In the case of a recreational water study undertaken by Soller et al. (2010a), the objective was to better understand the results from epidemiological studies conducted on the Great Lakes in the USA during 2003 and 2004 by identifying pathogens that could have caused the observed illnesses in those studies. Therefore, the selected health outcome for that study was gastrointestinal illness.

Table 5.2 Considerations in the selection of reference pathogens used in QMRA[a]

Reference pathogen	Epidemiological significance	Source quantification	Resistance to disinfection[b]	Persistence in the environment	Dose–response	Health outcomes
Bacteria						
Salmonella spp.	Common foodborne pathogen. High prevalence in environmental samples. Very diverse range of zoonotic sources, including avian (wild and domesticated), bovine, swine, domestic pets and reptiles.	Quantitative enumeration by culture.	Low resistance.	Can grow in the environment under suitable conditions (e.g. manure piles, specific food crops).	Published dose–response models fitted to clinical data for a range of *Salmonella* serotypes.	Usually self-limiting diarrhoea; sequelae include reactive arthritis and inflammatory bowel disease.
E. coli O157	Incidence is typically low. Most common zoonotic source is bovine; however, outbreaks linked to wild animals, including ducks and deer, have been documented.	Difficult to isolate. Few data sets available, and typically semiquantitative (presence/absence).	Low resistance.	Can persist under favourable conditions for long periods (weeks–months).	No clinical data. Published models fitted to outbreak data.	Probability of developing haemolytic uraemic syndrome is high, particularly in young (<5 years) children. Sequelae include end-stage renal disease.
Campylobacter jejuni	High incidence in developed countries. Most common zoonotic sources include avian (domestic and wild), bovine and ovine.	Difficult to isolate. Data sets are typically semiquantitative and not specific for *C. jejuni* species.	Low resistance.	Can persist in cold climates for long periods.	Published models fitted to both clinical and outbreak data.	Typically self-limiting diarrhoea. Sequelae include Guillain-Barré syndrome, reactive arthritis and inflammatory bowel disease.
Vibrio cholerae	Epidemic strains limited to several serotypes; common waterborne agent when endemic to a region.	Quantitative enumeration by culture.	Low resistance.	Persists in warm coastal waters and grows associated with zooplankton.	Published models fitted to clinical data.	Typically self-limiting diarrhoea, but if severe and no hydration treatment, can be lethal.
Viruses						
Adenovirus	Respiratory and gastroenteric types. Consistently present in human sewage. Prevalent in surface waters.	Culturable; enteric adenoviruses more difficult to culture.	Moderate resistance to free chlorine. Most resistant of viruses to low-pressure (254 nm) UV.[c]	Can persist in cold climates for long periods.	Model fitted to small clinical data set of HAdV-4 (respiratory) exposure via aerosols.	Respiratory illness, gastroenteritis, conjunctivitis, cystitis.
Enterovirus	Diverse range of viruses within enterovirus group (including coxsackieviruses). Relatively low prevalence in comparison with other enteric viruses.	Culturable.	Moderate resistance to free chlorine although most resistant of viruses to free chlorine.[d]	Can persist in cold climates for long periods.	Clinical data available for echovirus 12 and polioviruses; coxsackievirus B4 model fitted to animal (mice) data.	Mostly gastrointestinal illness, but also severe health outcomes, including meningitis, sepsis, myocarditis, poliomyelitis, link with type 1 diabetes.

Table 5.2 Considerations in the selection of reference pathogens used in QMRA[a] (continued)

Reference pathogen	Epidemiological significance	Source quantification	Resistance to disinfection[b]	Persistence in the environment	Dose–response	Health outcomes
Hepatitis A virus	Highly endemic in countries with limited water supply and sanitation; epidemic in other countries.	Culturable (but more difficult to culture).	Moderate resistance to free chlorine.	Long survival in water, resistant to high temperatures.	No published dose–response model (insufficient clinical data).	Hepatitis.
Rotavirus	Common cause of gastroenteritis in children and the elderly; particularly relevant in developing countries.	Culturable.	Moderate resistance to free chlorine.	Long survival in water.	Published dose–response model fitted to clinical data.	Important cause of death for children in developing countries.
Norovirus	Most common cause of gastroenteritis in developed countries. High person-to-person spread, and important for confined living conditions (e.g. cruise ships, nursing homes).	Not culturable. Only molecular data available.	Moderate resistance to free chlorine.	Appears to persist at cold temperatures for long periods.	Published dose–response model fitted to clinical data.	Usually self-limiting diarrhoea. Can lead to increased mortality in the elderly.
Protozoa						
Cryptosporidium parvum and *hominis*	Prevalent parasitic cause of gastroenteritis. *C. hominis* spread via humans, *C. parvum* also from zoonotic sources (cattle, sheep).	Standard methods have been widely applied, relatively large data sets available. Culturable methods available, but difficult to apply.	Very high resistance to free chlorine. Susceptible to UV.	Oocysts are resistant to environmental inactivation.	Published dose–response models fitted to clinical data for several strains.	Usually self-limiting diarrhoea that can last weeks. Can lead to long-term gastroenteritis in the immunocompromised.
Giardia lamblia	*G. lamblia*, *G. intestinalis* and *G. duodenalis* are synonyms for which assemblages A&B are human pathogens.	Standard methods have been widely applied, relatively large data sets available.	High resistance to free chlorine. Susceptible to UV.	Cysts less persistent than *Cryptosporidium* oocysts.	Published dose–response models fitted to clinical data.	Usually self-limiting diarrhoea that can last weeks. Recurrent diarrhoea episodes in some cases.

[a] Inclusions in the table are not exhaustive, but are intended to provide examples of the type of factors to be considered in the selection of reference pathogens for QMRA.

[b] Based on WHO (in preparation). Within pathogen species and groups there is likely to be variation in resistance, which could be further impacted by characteristics of the water supply and operating conditions. Resistance is based on 99% inactivation at 20 °C where, generally, low represents a Ct99 of <1 min.mg/L, moderate 1–30 min.mg/L and high >30 min.mg/L (where C = the concentration of free chlorine in mg/L and t = time in minutes) under the following conditions: the infective stage is freely suspended in water treated at conventional doses and contact times, and pH is between pH 7 and 8. For more detailed information on the comparative inactivation rates and susceptibility of pathogens to the full range of disinfectants and conditions, see WHO (2006a, b, in preparation) and associated supporting documents.

[c] See Hijnen, Beerendonk & Medema (2006).

[d] See Petterson & Stenström (2015).

6 EXPOSURE ASSESSMENT

The objective of the exposure assessment is to estimate the magnitude and frequency of exposure to each reference pathogen via the identified exposure pathways and hazardous events defined during problem formulation. This requires simplifying the environmental system to a defined, quantifiable exposure pathway.

In exposure assessment, the magnitude and frequency of exposure to each reference pathogen via each identified exposure pathway during each combination of event conditions are estimated. The frequency of exposure is defined in exposure assessment, but applied within risk characterization (Chapter 8) to quantify the overall probability of one or more infections/illnesses for a given time interval (e.g. per year), as a higher level of exposure may be tolerated for very rare exposure events, in comparison with daily exposure pathways. Exposure frequency can vary from daily (e.g. unboiled tap water [exposure within 24 hours is assumed to comprise a single event]) to a few days per week (e.g. certain food crops) to weekly (e.g. household garden watering with reclaimed effluent) or even sporadically (e.g. emptying a septic tank). The exposed population may vary from the vast majority (e.g. tap water) to a small subset (e.g. professional divers) of the total population.

As outlined in section 3.2, exposure assessment involves three steps: defining the exposure pathway(s); quantifying each component of the exposure pathway(s); and characterizing exposure, giving consideration to the units of dose (Box 6.1).

Box 6.1 Units of dose

The objective of the risk assessment is to combine the exposure dose with a dose–response model in order to estimate health risk outcomes. It is therefore important that the units of dose are consistent with the selected dose–response model. The majority of commonly applied dose–response models rely on the mean dose per exposure event, quantified as:

$$\text{Exposure dose} = C \cdot q \qquad \text{Eq. 6.1}$$

where:

C is the concentration of pathogens in the exposure medium; and

q is the amount of material ingested or inhaled per event.

The units of both C and q depend on the exposure pathway and may be based on mass for ingestion of soil or food (mg, g) and volume for water consumption (mL, L) or inhalation (m^3) per exposure.

The mean dose is a variable that may take any value greater than or equal to 0; however, microorganisms are discrete units, and the number of organisms to which an individual is exposed will always be a discrete number (i.e. 0, 1, 2, 3…). The microbial enumeration units (e.g. by culture or molecular methods; see section 6.2.1 and Annex C) are also an important consideration to test for harmony with those applied in the development of the dose–response model. This is discussed in more detail in section 7.1 and Annex D.

6.1 Defining the exposure pathway

The available data and the purpose of the risk assessment will drive how the exposure pathway is defined. Table 6.1 includes examples of different approaches for defining the exposure pathway. Any exposure pathway can be defined in terms of sources, controls and mechanisms of exposure (Fig. 6.1):

- **Sources:** The initial point of pathogen quantification. It may not always be possible or appropriate to quantify the concentration directly at the point of exposure. For example, for QMRA of drinking-water, the source could be the treated drinking-water (Borchardt et al., 2012), untreated source water (Signor et al., 2005) or even upstream faecal sources (Ferguson et al., 2007; Ferguson, Charles & Deere, 2009). For wastewater reuse scenarios, pathogens have been quantified in faecal sources (Ottoson & Stenström, 2003), raw sewage (Gale, 2005), treated effluent (Westrell et al., 2003) and irrigation water (Seidu et al., 2008).
- **Control measures (barriers):** Any environmental (e.g. residence time, sunlight, overland transport) (Muirhead, Collins & Bremer, 2006a,b; Sinton et al., 2007), engineered (drinking-water or wastewater treatment barriers) (Hijnen & Medema, 2010) or regulatory (e.g. withholding periods for crops) controls that are expected to lead to a loss or inactivation of pathogens (Hamilton et al., 2006) are identified and can be quantified. It is also important to note that recontamination from a secondary source (e.g. intrusion of groundwater to distribution networks; waterfowl on open ponds; human handling of food) may also occur.
- **Mechanisms of exposure (intake):** The pathway by which human exposure may result, which, depending on the context, may include intentional drinking (unboiled tap water) (Mons et al., 2007), unintentional ingestion (e.g. during recreational swimming) (Dufour et al., 2006; Suppes et al., 2014), aerosol ingestion (Schoen & Ashbolt, 2011) and food consumption (Hara-Kudo & Takatori, 2011).

It is helpful to construct a diagram identifying all system components that need to be quantified in the risk assessment.

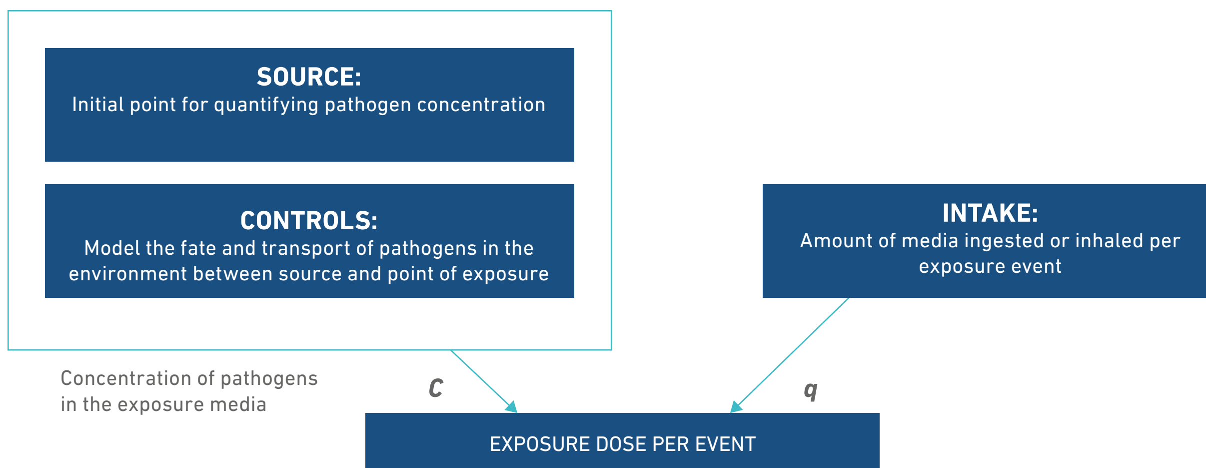


Fig. 6.1 Conceptual components for quantifying exposure (adapted from Petterson & Ashbolt, 2015)

As illustrated in Table 6.1, for a given exposure pathway, there may be several options for where to begin the quantification for the QMRA. The most straightforward option would be to quantify C, the concentration of pathogens in the exposure medium, and q, the amount of medium ingested per exposure event, directly. However, modelling pathogen reduction across barriers (rather than quantifying pathogen concentration directly in the exposure medium) may be preferred either 1) because of the limitations of enumeration data or methods (typically small and uncertain data sets, along with often large fluctuations in pathogen concentrations; e.g. Quintero-Betancourt, Peele & Rose, 2002; Eyles et al., 2003; Girones et al., 2010) or 2) to meet the purpose of the risk assessment (e.g. to evaluate different types of barriers or inform barrier management):

1) **Limitations with enumeration data:** Pathogen enumeration data may be unavailable for the exposure medium due to either resource constraints (financial or capacity) or methodological or sampling limitations.

   Pathogen concentration in the exposure medium may be too low to allow a consistent detectable concentration to be obtained. In the case of drinking-water, pathogen concentrations may be well below the limit of detection and yet still pose a risk to public health (Signor & Ashbolt, 2006; Smeets et al., 2007). Microbial risks associated with drinking-water exposures are therefore often assessed by quantifying the pathogen concentration in raw water and modelling reduction across treatment barriers (Teunis, Ever & Slob, 1999; Medema et al., 2006; Smeets et al., 2010).

Pathogen enumeration data may be too limited (small sample size) to describe the full range of fluctuations in the concentration. For example, pathogen monitoring data from surface water sources may exhibit a high proportion of zeros even when the water sources are known to be influenced by faecal sources. This is often due to the event-driven nature of microbial loading and the limitations of small monitoring data sets to capture these events. Modelling the pathogen concentration in faecal sources, followed by hydrologic modelling of contamination events, may therefore provide more useful information for QMRA than relying on monitoring data alone (Ferguson et al., 2007; Ashbolt et al., 2010; Sokolova et al., 2015); however, this approach can lead to overly conservative estimations of pathogen numbers.

2) **To meet the purpose of the risk assessment:** When the purpose of the assessment is to investigate processes and drivers associated with microbial risks, it is necessary to systematically model the environmental processes from source to exposure. Owing to the inherent uncertainties associated with quantifying exposure, the greatest value of QMRA may be not in the final quantification of risk, but rather in the exploration of system variables and risk drivers as a tool to support water safety management. In this case, each step within the system that needs to be explored must be included in the exposure pathway.

Therefore, defining an exposure pathway that fits with the available data and the purpose of the QMRA is critical and involves the following steps:

- **Describe the exposure pathway**, identifying the sources of pathogens and how they are transported from the source to the receptor.
- **Target the exposure pathway** in light of the available data and the purpose of the risk assessment. What scientific data are available? How would one expect the pathogen concentration to fluctuate throughout the system? Where is it possible and reasonable to measure the pathogen concentration? What factors and environmental processes should be considered within the risk assessment?
- **Define a conceptual model** that simplifies the exposure pathway as much as possible in order to allow quantification and to achieve the goal of the risk assessment. Define the scope of the exposure assessment by identifying the sources and processes/barriers that will be considered.

6.2 Quantifying each component of the exposure pathway

Once the exposure pathway is defined, each component needs to be quantified based on the best available scientific evidence. To ensure that the inputs are appropriate and representative, it is important to give consideration to the underlying nature of the scientific data, including site-specific data and data reported in the literature. This will include the applicability to the current setting, factors driving variability and uncertainty associated with the model and analytical methods, and the implications of the statistical analysis approach behind the reported values. Further detail regarding the interpretation and statistical modelling of microbial data is included in Annex C.

6.2.1 Source: concentration of pathogens in environmental media

In order to undertake the QMRA, the concentration of each reference pathogen (identified for the assessment during problem formulation; see Chapter 5) in the source material needs to be quantified. In many situations, pathogen enumeration data will be unavailable, and the assessment must be based on indicator data or expectations from the literature. Nevertheless, the following key principles relate to the evaluation of reference pathogen concentration in environmental media for QMRA:

- Quantitative concentration estimates need to draw on a systematic understanding of the individual site characteristics, including pathogen sources and events that may lead to fluctuations in pathogen concentration.
- All forms of scientific evidence that are available need to be brought together (sanitary survey, faecal indicators, microbial source tracking [MST] markers, pathogen enumeration data and information from the literature) to build an overall picture of the expected pathogen concentration in the environmental media for the individual site. Reliance on pathogen enumeration data alone, when other information is also available, is not recommended.
- Data on faecal indicator organisms (FIO) provide important and valuable information for risk assessment; however, the site-specific context and sources of FIO and pathogens need to be considered (e.g. role of sanitary survey and supplementary MST data) (see Annex C, section C4).
- The methods that were used for the enumeration of microorganisms must be known, including the implications for quantification, infectivity and viability (see Annex C, sections C1 and C3).

Table 6.1 Examples of different exposure pathways (sources and controls) for the same exposure scenario

Exposure scenario		Possible exposure pathways for quantifying *C*				
		Source	**Controls**			
Consumption of unboiled tap water from a conventionally treated municipal supply	Option 1	Finished drinking-water				Distribution network
	Option 2	Untreated raw water		Conventional treatment	Disinfection	Distribution network
	Option 3	Faecal sources	Dispersion and advection and die-off	Conventional treatment	Disinfection	Distribution network
Wastewater irrigation of food crops eaten raw	Option 1	Food crop at consumption				
	Option 2	Food crop at harvest			Processing and handling	
	Option 3	Irrigation water	Contamina-tion during irrigation	Inactivation in the field	Processing and handling	
Swimming in a freshwater lake	Option 1	Lake water				
	Option 2	Faecal sources	Dispersion and advection	Inactivation		

Pathogen enumeration considerations	Purpose of the risk assessment
• Concentration is often too low for detection • Exposure is often linked to fluctuations in source water quality or treatment failures, which are difficult to characterize with small data sets	• To evaluate the level of risk to the consumer • To evaluate whether additional treatment is required
• Concentration is higher in surface water and hence more detects • Fluctuations in surface water concentration may mean that capturing the full range of variability is not possible	• To evaluate the level of risk to the consumer • To evaluate whether additional treatment is required • To investigate the robustness of treatment barriers and the impact of failure events
• Fluctuations in pathogen concentration in the faecal sources can be characterized • Pathogen loading events can be modelled • Site-specific pathogen concentrations may be difficult to quantify accurately • Modelling the combined influence of different human and non-human sources is a challenge	• To identify the influence of pathogen loading events and the importance of the epidemiology of incidence of infection on surface water quality
• Fluctuations in concentration are difficult to detect with small sample sizes	• To evaluate the level of risk to the consumer associated with the food crop
• Quantification (and the need for treatment/processing) of highly contaminated crops is possible	• To evaluate the level of risk to the consumer associated with the food crop • To separate the influence of human handling on potential risk to consumer
• Higher and more consistent concentrations of pathogens in wastewater can be quantified	• To evaluate the level of risk to the consumer • To separate the influence of human handling on crop contamination • To evaluate the effect of irrigation practices on water quality and exposure • To calculate the regulatory crop withholding period given expected inactivation under different climatic conditions
• Fluctuations in lake water concentration may mean that capturing the full range of variability is not possible	• To evaluate the level of risk associated with swimming in surface water
• Higher and more consistent concentration of pathogens in faecal sources can be quantified • Site-specific pathogen concentrations may be difficult to quantify accurately • Modelling the combined influence of different human and non-human sources is a challenge	• To identify and prioritize high-risk events • To relate measured faecal indicator concentration to pathogen concentrations and hence risks • To separate the influence of other swimmers on water quality

- The steps involved in the processing of the environmental samples and the implications for quantifying method recovery must be understood (see Annex C, section C1).
- The uncertainty in the quantified concentration must be explicitly considered; how representative are the data for the pathogens of interest? What conditions are *not* represented by the data, and are these conditions important for the QMRA?
- Quantified units of pathogens must be comparable with the units used in the dose–response assessment.
- Although statistical simplifications are often necessary (and appropriate for a screening-level or tier 1 analysis), the implications of these simplifications need to be considered (see examples in Annex C).

6.2.1.1 Variability in pathogen concentration

The true, underlying concentrations of pathogens and indicators in environmental media will vary both temporally (with time) and spatially (between sites) because of many factors summarized in Table 6.2. Several of these listed factors can exhibit considerable fluctuations and may be event driven; examples of events or incidents that may lead to significant fluctuations in pathogen concentration are also provided in Table 6.2.

Each of these factors will influence the magnitude of and variability in the pathogen and indicator concentrations at a particular site, and hence generalizing pathogen concentrations in environmental media is difficult. A summary of indicative pathogen concentration ranges in environmental media (i.e. water and sewage) is included in each of the guidelines: GDWQ Table 7.6 (WHO, in preparation); GWEG Table 3.2 (WHO, 2006a); and GREC Table 4.1 (WHO, 2003). Note that the reported concentrations span several orders of magnitude as a result of true variability in concentration, differences in contamination levels between locations and systems, and the influence of method uncertainties (see section C1 in Annex C). Given the considerable differences between studies, site-specific information is very important for reducing uncertainty in quantifying pathogen concentrations.

6.2.1.2 Site-specific evaluation of pathogen sources

The first stage in gathering information for quantifying pathogen and indicator concentrations is a site-specific evaluation of pathogen sources and their variability. Some of this information should be collected as part of the system description and hazard identification steps of the safety plan process. The objective is to create a quantitative expectation of the magnitude of and variability in pathogen concentration by undertaking the following steps, in essence undertaking a sanitary survey from source(s) to point(s) of exposure:

- **Identify faecal sources to the environmental media:** For wastewater use applications, this may be straightforward (sewage or treated wastewater); for surface water sources used for drinking-water or for recreational waters, a watershed-based sanitary survey of faecal sources associated with surface and subsurface water hydrodynamic models can be used to collect this information.
- **Identify factors driving the pathogen concentration in the faecal sources (including events):** How the pathogen concentration within identified faecal sources is expected to vary should be determined. The impacts of the important factors driving pathogen concentration for each exposure scenario should be identified and estimated (Table 6.1). Disease surveillance records may provide insights regarding prevalence rates (and seasonal fluctuations) for the specific region.
- **Identify factors driving the contribution of the faecal sources to the point of quantification (including events):** The factors leading to transport of pathogens from their source to the point of exposure (quantification) should be identified and evaluated.
- **Compare the study site with published information in the literature or through modelling:** Either through direct review of published results from other similar sites or through numerical modelling of faecal sources, an understanding of reference pathogen concentrations in comparable environmental media should be developed.

Table 6.2 Factors and associated events influencing the magnitude and variability of pathogen concentrations in environmental media

Environmental media	Factors influencing magnitude and variability in pathogen concentrations	Events or incidents that may lead to significant fluctuations in pathogen concentrations
Faecal samples and non- or low-water waste sources such as latrines and pour-flush toilets (combined faecal material from a population[a])	• Incidence of infection in the population: only infected (colonized) individuals are assumed to excrete pathogens, whereas faecal indicators are generally commensals and so always present. • Number of individuals contributing to the source: depending on the incidence of infection, the number of individuals contributing is a strong driver of variability in pathogen and indicator densities. • Variability in excretion density (counts of pathogens or indicators in the faeces of a colonized individual), both between individuals and in a single individual over the course of an infection. • Impacts of animal faecal waste sources, including types and numbers of animals and their accessibility to exposure media, such as drinking-water sources.	• Outbreak of the targeted pathogen in contributing population
Raw wastewater	All of the above plus.... • Household water usage rates • Extra flows and infiltration to the sewer (e.g. industrial wastewater, rainfall-induced surcharges) • Pathways and magnitudes of faecal waste and sewage inputs to exposure sources	The above plus.... • Major rainfall events
Treated wastewater	All of the above plus.... • Wastewater treatment process efficacy	All of the above plus.... • Suboptimal treatment • Process failure?
Surface waters	All of the above plus.... • The combined influence of human and non-human faecal sources • Climatic/environmental factors influencing pathogen and indicator persistence • Attenuation and transport from source to sampling location	All of the above plus.... • Rainfall/runoff events • Reservoir short-circuiting • Seasonal decreases in water quantity and quality due to evaporation and infiltration
Groundwater	All of the above plus.... • Sewage exfiltration, septic seepage, manure seepage and infiltration of surface water and subsurface flow to aquifer	All of the above plus.... • Fractured aquitard or highly porous soil • Higher rate of well extraction • Seasonally high water tables and soil saturation with water due to rainfall or other conditions
Drinking-water	All of the above plus.... • Drinking-water treatment barrier efficacy • Post-treatment recontamination (distribution, household storage)	All of the above plus.... • Treatment interruption and suboptimal treatment • Process failure • Contamination in distribution network (e.g. intrusion, backflow) • Household contamination

[a] A faecal sample from a single individual would be expected to exhibit different characteristic variability compared with that from a large human wastewater source. Pathogens would be expected to be present only if the individual is infected; concentration as excreted would depend on the severity and stage of infection; and likelihood is based on a particular pathogen's prevalence in a community, which may well exhibit (seasonal) fluctuations for a specific pathogen.

The site-specific pathogen data can be checked against the sanitary survey, indicator data and pathogen data from published studies in comparable water systems. When monitoring results provide consistent values between indicators and pathogens and with expectations, including when comparing against the published literature, there is added confidence in the pathogen estimates; alternatively, when the monitoring results are inconsistent with the expectations (e.g. no detection of pathogens from a highly faecally impacted catchment or high results with no known pathogen sources), the need for further investigation is triggered.

6.2.1.3 Microbiological monitoring data

Following the initial site assessment, and provided they are available, the monitoring (routine and/or research) data need to be interpreted and analysed. Data sets from microbiological monitoring programmes may include results for FIO and/or human pathogens:

- **Faecal indicator organisms (FIO):** FIO, such as *E. coli,* are usually not human pathogens, but rather are selected for enumeration owing to their ease of analysis (relative to pathogens) and their ability to represent the potential presence of pathogens in faecally contaminated environmental media, among other factors (e.g. see section 7.4 of GDWQ; WHO, in preparation). The relationship between pathogens and FIO depends on how specifically faecal in origin the particular indicator is, how host specific it is and how environmentally persistent it is in comparison with the targeted human pathogens in a QMRA. Interpretation of indicator data therefore requires careful consideration of likely faecal sources from the sanitary survey data. A full discussion of FIO and their application for predicting pathogen concentrations for QMRA is included in section C4 in Annex C.
- **Pathogen data:** Analysis of human pathogens from environmental samples is complex, prone to highly variable analytical results and costly. Understanding of the methods associated with the processing, isolation and identification of pathogens is extremely important for quantitative interpretation of the results for QMRA (see Annex C). Important factors for consideration include:
  - How many samples were collected, and how many individual observations are there per sample?
  - Were grab samples collected at one point in time, or were composite samples collected incrementally over a duration of time (e.g. 24 hours)?
  - What characteristics of the target organism were identified, and how specific are these characteristics for the reference pathogen of interest (see Annex C, section C1)?
  - What was the quantification principle of the method (e.g. most probable number [MPN], colony-forming units [CFU], plaque-forming units [PFU], copies per volume unit)? What is the uncertainty associated with the quantified value (see Annex C, section C2)?
  - How specific was the method for targeting pathogens infectious to humans (see Annex C, sections C1 and C3.3)?
  - What was the recovery efficiency of the method? Were there quantitative controls for method recovery efficiency (see Annex C, section C3.2)?
  - What do the results mean regarding the infectivity of the identified organisms (see Annex C, sections C1 and C3.3)?

Unless there is a local capacity for pathogen analysis in water and/or wastewater using the preferred analytical methods for which sample sizes are larger, the capacity for analysis of faecal indicator microbes, typically *E. coli*/enterococci bacteria, is often greater, and larger numbers of samples can often be analysed to better characterize central tendencies and variability.

6.2.1.4 Statistical descriptors and statistical models

Inferring quantitative values for the exposure assessment from the monitoring data requires consideration of the representativeness of the sample and the most appropriate descriptors to be estimated:

- **Choosing a representative sample:** For waterborne pathogens, sample numbers are often small (<20) and sometimes very small (<5). Consequently, a critical question for analysis of any microbial data set is: what conditions do the data represent? Review of the sampling regime in comparison with the site-specific evaluation of variability (discussed above) provides information regarding what conditions are represented by the samples. It is important to know if the samples were collected randomly (e.g. following a defined procedure that is independent of environmental conditions, such as on the first Monday of a month), in response to a perceived risk (hazardous event) or under conditions where the pathogen loads during sampling were expected to be

particularly high (e.g. following rainfall/snowmelt event or a sewer bypass event), relatively low (e.g. during low-flow conditions) or perhaps average.

- **Choose which descriptors or parameters are to be estimated:** The value to be estimated depends on the purpose of the assessment (Chapter 5), the exposure scenarios and how the risk will be characterized (Chapter 8). Point estimates may include the mean, median, 95th quantile or maximum (i.e. worst-case) pathogen concentration. Alternatively, a parametric probability distribution may be fitted to the data to describe the variability in concentration (see Annex B and section C3 in Annex C for more examples). Considerations in the selection of statistical descriptors or fitting of parametric distributions include the following:
  - Do the monitoring data represent the original observations (see Annex C, section C2)?
  - What are the characteristics of the data? The data may be qualitative (presence or absence in a given volume); discrete (only contains integers [0, 1, 2, 3...]); categorical (can hold one of a range of distinct values, as in the case of MPNs); or continuous (can hold any value within a set range, such as in the case of reported concentrations) (see Annex C, section C1.3).
  - What should be done about zeros? Microbial count data often include (sometimes many) zeros (see Annex C, section C3.1).
  - What should be done about censored data? Reported microbial concentrations typically contain left-censored (less than) and right-censored (greater than or too numerous to count [TNTC]) data (see Annex C, section C3.1).
  - How should MPNs be analysed? Reported concentrations based on MPN inference are predicted concentrations from presence/absence data and frequently include left- and right-censored outcomes (see Annex C, sections C3.1 and C3.4).
  - How should the concentrations be corrected for imperfect recovery? Only a fraction of the microorganisms originally present will be detected by the analytical method (see Annex C, section C3.2).
  - How can the concentration be corrected to describe only the human-infectious fraction? Depending on the method, only a fraction of the total number of microorganisms detected will be infectious to humans (see Annex C, section C3.3).

6.2.1.5 Interpretation of published data

In many cases, site-specific data are not available for quantifying pathogen concentration in source material, and the assessment must rely on data published in the literature. Firstly, it is important to evaluate how similar the site from the published study is to the specific study system. All of the factors that drive the true variability in pathogen concentration in environmental media (listed above) need to be considered, including disease incidence, sanitation, climate and topography. Differences need to be documented so that the potential impact on the risk results can be assessed.

Secondly, when drawing on data from the literature, all of the previously mentioned factors related to methodology and data reporting (listed above) need to be considered. If the specific details are not reported with the published data, it is necessary to:

- evaluate the magnitude of the unknowns and whether the data are suitable for application in the QMRA;
- make a conservative assumption in the absence of specific information; and
- document the assumption and evaluate its importance on the study outcomes in risk characterization.

6.2.2 Quantifying efficacy of control measures

The second step in quantifying each component of the exposure pathway is to evaluate the performance of each identified control measure in reducing the concentration of reference pathogens. Depending on the purpose of the risk assessment, consideration of the variability in performance and the influence of events may be critical to the analysis. In quantifying reductions in pathogens as a result of control measures, several issues must be carefully considered. These include the following:

- Quantitative control reduction estimates need to draw on a systematic understanding of the individual barrier, including mechanisms of pathogen reduction and events that may lead to fluctuations in performance.
- Engineered barriers are usually operated and controlled; how do operation and control affect barrier performance and variability?

- All forms of scientific evidence that are available need to be brought together (process and hydrological conditions, surrogate data, online monitoring data and information from the literature) to build an overall picture of the expected pathogen reduction efficacy for the individual site.
- An empirical approach to quantifying reduction efficacy may not be appropriate for all control measures, and the processes and mechanisms of the barrier need to be considered in the quantification.
- Application of data from the literature requires careful consideration regarding the representativeness of empirical-based barrier quantification from another site; and the representativeness of parameter values (e.g. inactivation coefficient) for mechanistic models.
- Although statistical simplifications are often necessary (and appropriate for a screening-level or tier 1 analysis), the implications of these simplifications need to be considered.

The uncertainty in the quantified concentration must be explicitly considered: How representative are the data for the pathogens of interest? What conditions are *not* represented by the data, and are these conditions important for the QMRA?

Microbial reduction in QMRA is most often quantified in terms of $\log_{10}$ reduction (1 $\log_{10}$ = 90%; 2 $\log_{10}$ = 99%; 3 $\log_{10}$ = 99.9%; etc.). Compilations of information on barrier reduction have been included in the relevant guidelines: drinking-water treatment process technologies for large systems (GDWQ Table 7.7; WHO, in preparation) and households (GDWQ Table 7.8; WHO, in preparation) and Table A2.4 in WHO (2011); wastewater treatment technologies (GWEG Table 5.2; WHO, 2006a); and wastewater health protection measures (GWEG Table 4.3; WHO, 2006a). In addition, for drinking-water systems, supporting documentation is available on treatment barrier performance (LeChevallier & Au, 2004; Hijnen & Medema, 2010). These reviews provide a starting point for estimating pathogen reduction efficacy for QMRA; however, there is considerable uncertainty associated with applying these values to a specific system, as the reported efficiency of barriers in the scientific literature is highly variable. For example, in the review and meta-analysis undertaken by Hijnen & Medema (2010), the mean reported $\log_{10}$ reduction of viruses by conventional drinking-water treatment (coagulation, flocculation, rapid granular filtration) across seven published studies varied from 1.2 to 5.3 $\log_{10}$; they calculated a mean elimination capacity, based on their meta-analysis of studies, of 3 $\log_{10}$. Reference to the GDWQ (Table 7.7; WHO, in preparation) indicates that the removal of viruses via coagulation, flocculation and sedimentation plus granular high-rate filtration would be expected to vary between 0.1 and 6.9 $\log_{10}$. This is a very broad range and would lead to very high uncertainty in the calculated risk; site-specific information on removal performance would be necessary to reduce this uncertainty, in essence understanding how well the treatment process is designed, operated and controlled (through validation and monitoring) and the "good" and "bad" days in treatment at the specific site (Gale, 2002).

6.2.2.1 Variability in barrier efficiency

Barrier efficacy will vary both spatially (between sites) and temporally (over time between sampling occasions). Whereas different types of barriers are grouped together due to their similarity – for example, "conventional drinking-water treatment", "activated sludge" or "waste stabilization ponds" – there are often considerable differences between study sites, including scale (pilot scale/full scale; capacity), details of design and operation (geometry, media, loading rates, hydrodynamics), water quality characteristics (organic content, pretreatment, temperature) and climate (temperature, solar radiation, precipitation), that will affect the microbial reduction performance. Although it may be desirable to generically characterize barriers, variability within categories is to be expected.

For any given barrier, efficacy will also be expected to vary over time. Barrier performance can be dependent on climatic factors, hydraulic loading, hydrodynamics and water quality characteristics. Therefore, seasonal and specific event factors (e.g. rainfall, upstream loading) will influence performance. In addition, depending on the barrier, performance can also be affected by system management decisions or failures (e.g. backwash frequency, chemical dosage rates). A summary of the factors and associated events influencing the variability in treatment performance for some barriers is given in Table 6.3. In addition to these factors, treatment interruptions caused by periodic lack of electricity can lead to pressure changes and potential ingress of contamination within the distribution system (Krumpel & Nelson, 2013).

Table 6.3 Examples of factors and associated events influencing the magnitude and variability of pathogen reduction for some common barriers

Barrier	Factors influencing magnitude of and variability in removal	Events or incidents that may lead to significant fluctuations in pathogen reduction
Engineered treatment barriers		
Coagulation/flocculation/ sedimentation	• Coagulation efficacy (including pH, alkalinity, dosing)	• Coagulant dosing failure • Rapid change in water quality
Rapid granular filtration	• Loading rate • Backwash cycle	• Breakthrough
Slow sand filtration	• Residence time	• Removal of schmutzdecke
Biological treatment	• Retention time • Water temperature	• Overloading during rainfall events
Microfiltration/ultrafiltration	• Pore size • Loading rate • Surface characteristics of membrane	• Membrane breakage • Membrane fouling • Failure of membrane seal, leading to water bypass of membrane
Disinfection	• Disinfectant dose • Contact time	• Dosage failure
Waste stabilization ponds and wetlands	• Retention time • Secondary contamination by waterfowl; may lead to an increase in some pathogens	• Rainfall event high flow • Outbreak of human-infectious strains in avian population
Point-of-use drinking-water treatment devices	• As above (depending on treatment method)	• Non-usage • Faecal contamination by handling and in storage
Distribution network	• Pipe length, pressure, integrity	• Biofilm sloughing • Intrusion
Environmental barriers		
Overland transport	• Method of deposition • Topography and soil type • Overland flow	• Large rainfall runoff event
Reservoirs and ponds	• Retention time • Secondary contamination by waterfowl, livestock and feral animals may lead to an increase in pathogens	• Large rainfall runoff event • Short-circuiting (stratification) • Outbreak of human-infectious strains in avian population
Biosolid storage	• Time • Environmental conditions	• Regrowth of pathogenic bacteria
Inactivation on food crops	• Environmental conditions • Pathogen protection on crops	
Regulatory barriers		
Prevention of access	• Elimination of exposure	• Breach of regulation
Buffer zones from irrigation	• Aerosol transport	• Unusual climatic/wind conditions

6.2.2.2 Site-specific evaluation of pathogen removal

In order to quantify the pathogen removal efficiency of a particular treatment barrier, it is necessary to understand the pathogen removal mechanisms and how the performance of those mechanisms is likely to vary for a specific site. The following steps should be performed:

1) **Identify the mechanisms of pathogen removal.** Why does the barrier result in a reduction in pathogen concentration? Some barriers physically remove pathogens from the water column – for example, sedimentation tanks, granular filtration, microfiltration and ultrafiltration. Barriers may also rely on inactivation either due to exposure to environmental conditions alone (sunlight, temperature, desiccation) or in the presence of a disinfectant.
2) **Identify the factors that drive the effectiveness of the removal mechanism.** Physical removal by filtration depends on the medium, pore size and loading rates and may also depend on effective chemical coagulation. Barriers dependent on inactivation require that sufficient residence time is achieved under adverse conditions or that a sufficient dose of disinfectant is delivered. What are the expected conditions for the study site?
3) **Identify the circumstances under which those dependent factors might fluctuate.** For filtration, loading rates and effective pore size can vary over the filter cycle; failure to deliver appropriate coagulant dose may lead to poor coagulation and hence poor removal; residence time of wetlands, ponds and reservoirs can be significantly affected by rainfall events, leading to short-circuiting; and seasonal conditions may support pathogen persistence, such as cooler temperatures and shorter days in winter. How might the expected conditions vary due to short-term (process fluctuations and failures, rainfall) and medium-term (seasonal) events?
4) **Compare the study site with published information in the literature.** There are many reviews available in the literature on a wide range of barriers. How is the study site expected to compare with reported ranges in the literature?

Site-specific evaluation not only provides an important first indication of how the barrier performs, but also informs future data collection and analysis. The evaluation will inform further investigations and which surrogate organisms are most suitable to describe the mechanism and pathogen of interest.

There are two main approaches to quantify barrier efficiency for QMRA:

1) **Empirical approach:** An empirical approach to quantify reduction performance is data driven and relies on observations of pathogens or surrogates before and after the barrier. The reduction efficiency is calculated as the difference in concentration (usually on a $\log_{10}$ scale). More detail is presented in Annex C, section C5.
2) **Mechanistic approach:** Rather than quantifying a single set of experimental conditions, the objective of the mechanistic approach is to characterize specific factors that drive the removal mechanism (e.g. residence time of a reservoir; persistence of pathogens during disinfection; hydraulic characterization of disinfection chambers). Log_{10} reduction is modelled based on the quantification of these driving factors. More detail is presented in Annex C, section C6.

The selected approach depends on the nature of the barrier, the scientific evidence and site-specific information available for quantification. The empirical approach is best suited for barriers that operate close to or in steady state (stable loading conditions) and that exhibit limited variation in operational conditions. If the process is dynamic (constantly responding to changes in loading or operational conditions) or known to be driven by measurable factors (e.g. hydraulic residence time), then a mechanistic approach may be more useful for quantification.

6.2.2.3 Aggregating multiple barriers for water/wastewater treatment performance

Water and wastewater treatment trains often consist of multiple barriers; indeed, a multiple-barrier approach is recommended to support water safety (WHO, in preparation). When combining multiple treatment barriers in series, it is common practice in QMRA to sum individual unit $\log_{10}$ credits to estimate the overall removal efficacy. In reality, however, treatment units do not act independently, but are designed to work in combination, such as the removal of suspended matter in coagulation/filtration processes, which makes chlorination more effective. Under nominal conditions, this means that the overall removal is at least as good as the sum of the removal of the unit processes, and may even be better, owing to synergism between processes. However, performance of treatment processes can also be negatively impaired by deterioration of water quality or performance of upstream barriers. For example, an increase in the organic content of surface water caused by a rainfall event may change the coagulation conditions of conventional drinking-water treatment, leading to poorer removal performance and higher concentration of particulates in the effluent. These particulates may protect

pathogens during disinfection, and hence the performance of the disinfection process may be compromised. QMRA should aim to incorporate the interaction between treatment processes; however, there are limited examples in the literature to demonstrate how this could be done in practice. According to current practice, a conservative estimate of treatment performance may be selected in order to account for these potential process interactions; however, this practice accounts only for the situation where the overall $\log_{10}$ reduction is less than the sum of the components.

6.2.3 Intake: exposure volumes and frequencies

The final stage in the exposure assessment is to quantify the magnitude and frequency of exposure for a range of different activities, including voluntary consumption (e.g. drinking-water, food crops) and involuntary consumption (e.g. accidental ingestion while swimming, aerosol ingestion/inhalation, dermal contact). Two approaches can be taken in QMRA to quantify exposure:

1) quantify exposure frequency and magnitude based on scientific data; and
2) select reference values for exposure magnitude and frequency.

The selection should be based on the availability of scientific data and the purpose of the risk assessment. Reported data and selected values for QMRA are summarized in the following sections.

6.2.3.1 Drinking-water

Drinking-water consumption varies between cultures, climates and individuals. For example, in the review of drinking-water consumption data by Mons et al. (2007), they reported that the mean daily consumption of cold tap water varied between 0.10 and 1.55 L across different regions (a subset of the data summarized in that review is included in Table 6.4). In their review, Mons and coworkers (2007) identified many important aspects associated with the analysis and interpretation of consumption data, including the following:

- The way in which the data were collected influenced the estimated consumption volume. Estimations of drinking-water consumption were higher in the questionnaires than in the diaries.
- The fraction of the population that did not drink tap water (non-consumers) varied widely between studies and needs to be considered.
- Distributions of drinking-water consumption were generally left skewed, and therefore the mean (being higher than the median) was considered a suitable conservative point estimator for QMRA.
- When defining a distribution of consumption, whereas the lognormal distribution has often been used to describe daily consumption volume, tap water consumption data are typically discrete in nature (number of glasses consumed per day); therefore, the Poisson distribution was considered more appropriate and, when non-consumers were accounted for, provided a suitable fit to the reviewed data sets.
- When no country-specific data are available, Mons and coworkers (2007) recommended the application of data from the Melbourne (Australia) diary study (Poisson, $\lambda = 3.49$ glasses·d^{-1}) as a conservative input, as the consumption level for this study was relatively high.

Despite the availability of drinking-water consumption data, it is common in QMRA to use a reference value for exposure – for example, many drinking-water QMRAs have relied on an assumption of 2 L per person per day (Regli et al., 1991; Rose, Haas & Regli, 1991; Asano et al., 1992) or 1 L per person per day (Masago et al., 2004; Howard, Pedley & Tibatemwa, 2006; WHO, in preparation). The advantage of a reference value is that it is transparent in the magnitude of exposure that is being represented by the risk calculations, which can be helpful, particularly when the true magnitude of exposure for the particular population is unknown.

Table 6.4 Summary of drinking-water consumption data[a]

Country	Study type	*N*	Mean consumption (L)					Reference
			Cold tap water	Heated tap water[b]	Total tap water	Bottled water	Total water	
USA	24 h recall	15 303	0.508	–	0.927	0.161 0.737[c]	1.232 1.241	USEPA (2000)
USA	Q/D	26 081	–	–	1.108	–	1.785	Roseberry & Burmaster (1992)
USA	Q	1 183	–	–	1.91	–	–	Williams, Florez & Pettygrove (2001)
Canada	D	970	–	–	1.34	–	–	EHD (1981)
Netherlands	FFQ	3 200	–	–	1.5	–	–	Foekema & Engelsma (2001)
Netherlands	Q	4 620	0.25 0.38[c]	–	–	–	1.14	Haring et al. (1979)
Netherlands	Q	–	0.153	–	–	–	–	Teunis et al. (1997)
Netherlands	D	6 250	0.178	–	–	–	–	Anonymous (1998); Hulshof, personal communication, 2003
Denmark	FFQ	195	0.5	1.08	–	–	1.58	Dangendorf (2003), F. Dangendorf, personal communication, 2004
France	D	373 (w) 427 (s)	0.77[d] 0.90[d]	0.54[c] (w) 0.61[c] (s)	1.55[c] (w) 1.78[c] (s)	0.85[c] (w) 1.07[c] (s)	1.83[c] (w) 2.19[c] (s)	Gofti-Laroche et al. (2001)
Spain	Q	157	0.86	0.94	1.80	0.06	1.86	Westrell (2004)
Spain	Q/D	35 40	1.14 1.55	0.81 1.05	1.95 2.58	–	–	Berg & Viberg (2003)
United Kingdom	D	3 564	0.103 0.203	0.785 1.065[c]	0.955 0.958	–	–	Hopkin & Ellis (1980)
United Kingdom	D	1 018	–	–	1.138	–	–	DWI (1996)
United Kingdom	Q	416 421	0.704 1.187	–	–	–	–	Hunter et al. (2004)
Australia	Q D Q	253 234 231	0.991 0.892 0.964	–	–	–	–	Robertson, Sinclair & Forbes (2000); Robertson et al. (2000); M. Sinclair, unpublished data
Australia	Q (Melbourne)	950	0.842	–	–	–	–	Robertson, Sinclair & Forbes (2000); Robertson et al. (2000); M. Sinclair, unpublished data

Table 6.4 Summary of drinking-water consumption data[a] (continued)

Country	Study type	*N*	Mean consumption (L)					Reference
			Cold tap water	Heated tap water[b]	Total tap water	Bottled water	Total water	
Australia	Q (Adel-aide)	644	0.718	–	–	–	–	Robertson, Sinclair & Forbes (2000); Robertson et al. (2000); M. Sinclair, unpublished data
Japan	Q	1 188 (w) 1 278 (su)	0.255 (w) 0.542 (su)	0.870 (w) 0.617 (su)	1.125 (w) 1.159 (su)	0.513[e] (w) 0.777[e] (su)	1.638 (w) 1.936 (su)	Ohno, Asami & Matsui (2013)

D: diary; FFQ: food frequency questionnaire; Q: questionnaire; s: spring; su: summer; w: winter

[a] If consumers and non-consumers were considered separately, data are presented for the total population, including non-consumers.

[b] The temperature of hot drinks is typically not known by the consumer and cannot be assumed to have reached the boiling point; hence, the term “heated” is used.

[c] Data for consumers; non-consumers not included.

[d] Cold tap water consumed at home directly from the tap (cold tap water added to, for example, lemonade and cold tap water consumed outside the house are not included in this figure).

[e] Bottled water including commercial beverages.

Source: modified subset (average consumers only) of Table 1 in Mons et al. (2007)

6.2.3.2 Wastewater reuse

There are a few examples of QMRA studies that have sought to draw on scientific data to support input assumptions, such as the application of the upper limit of daily soil ingestion for children for sludge application studies (Westrell et al., 2004); vegetable consumption data (Gale, 2005; Hamilton et al., 2006); and plume modelling for aerosol dispersion of wastewater used for spray irrigation of crops (Höglund, Stenstroöm & Ashbolt, 2002). The most common approach for wastewater reuse QMRA is to use reference values for exposure. In the guidelines, examples of reference values include 1–10 mg of soil ingested per person per day for 100 days per year for workers involved with highly mechanized agriculture (Table 3.14 GWEG); 10–100 mg soil ingested per person per day for 300 days per year for labour-intensive agricultural workers (Table 3.15 GWEG); 100 g of wastewater-irrigated lettuce consumed per person every 2 days (Table 3.17 GWEG); and 100 g of onions consumed per person once per week for 5 months (Table 3.18 GWEG) (WHO, 2006a). These values are assumed to conservatively represent the exposure scenario being modelled. For consumption of food crops, many countries have national statistics that may be useful. Further examples of reference values for wastewater reuse are included in Case-study 5 in Annex A.

6.2.3.3 Recreational water

For recreational studies involving primary (swimming, bathing) contact, Dufour and coworkers (2006) investigated the volume of water consumed by pool swimmers, providing observational data that have subsequently been used for QMRA (Schoen & Ashbolt, 2010; Soller et al., 2010b). Fifty-three recreational swimmers participated using a community swimming pool disinfected with cyanuric acid–stabilized chlorine. The swimmers were asked to actively swim for at least 45 minutes and to collect their urine for the next 24 hours. The predicted median volume of water swallowed by all participants combined was approximately 19 mL. Results of the study indicated that non-adults ingest slightly more water than adults during swimming activity (Fig. 6.2). The average amounts of water swallowed by non-adults and adults were 37 mL and 16 mL, respectively. There is some evidence to suggest that risks may be different for surfers (Stone et al., 2008; Tseng & Jiang, 2012).

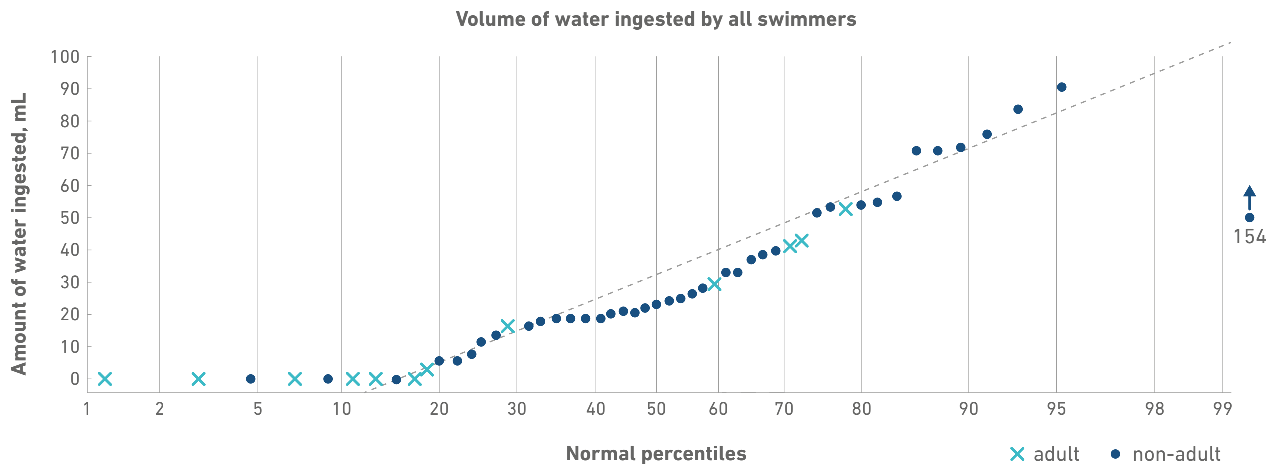


Fig. 6.2 Ingestion volumes during recreational activities (Dufour et al., 2006)

A range of reference exposure values has also been applied in QMRA studies, from 10 mL to represent accidental gulping (Steyn, Jagals & Genthe, 2004) up to 100 mL to represent unintentional ingestion during full immersion activities (Asano et al., 1992; Mena et al., 2003; Steyn, Jagals & Genthe, 2004; Diallo et al., 2008) (see Table 6.5).

Table 6.5 Reference values for unintentional consumption during recreational activities

Risk assessment study	Location	Exposure pathway	Input to QMRA	Reference for input
Asano et al. (1992)	California, USA	Swimming	100 mL	Haas (1983)
Mena et al. (2003)	Not specific	Swimming	100 mL: 1, 5 and 10 days of exposure	NG
Craig, Fallowfield & Cromar (2003)	Australia, coastal	Swimming	Uniform 20–50 mL	Ashbolt, Reidy & Haas (1997)
Steyn, Jagals & Genthe (2004)	South Africa	Full immersion	100 mL	Genthe & Rodda (1999); Haas, Rose & Gerba (1999)
Steyn, Jagals & Genthe (2004)	South Africa	Intermediate	50 mL	Medema et al. (2001)
Steyn, Jagals & Genthe (2004)	South Africa	Other – accidental gulping	10 mL	Genthe & Rodda (1999); Medema et al. (2001)
Westrell et al. (2004)	Sweden	Swimming	50 mL 50 times per year	NG
van Heerden et al. (2005)	South Africa	Swimming	30 mL	Crabtree et al. (1997)
Diallo et al. (2008)	Thailand	Swimming	100 (50) mL	NG

NG: not given

Secondary contact recreational activities (boating, fishing, etc.) can also lead to unintentional ingestion of water. The likelihood and volume of consumption are assumed to be lower in comparison with primary contact; however, there are limited quantitative data available. Dorevitch et al. (2011) estimated the volume of water consumption for a range of limited-contact recreational activities, including canoeing, kayaking and fishing, based on self-reporting of water consumption by category (“drop”, “teaspoon”, “mouthful”) among 2705 participants. Mean and upper confidence limit estimates of water ingestion were about 3–4 mL and 10–15 mL, respectively. The frequency of swallowing at least a teaspoon of water during limited-contact recreation was around 1% of study participants.

7 HEALTH EFFECTS ASSESSMENT

Dose–response relationships describing the relationship between exposure and the probability of infection or illness are identified and evaluated for each reference pathogen. The likelihood and consequence of various health outcomes (potentially including infection, illness and/or long-term health impacts) are assessed.

7.1 Dose–response relationships

Most risk assessments use dose–response information from the published literature. Care must be taken that the selected model and parameter values are appropriate for the risk study and that when interpreting the calculations, the implications of the assumptions underlying the selected numbers are considered. The following key principles need to be considered for application of dose–response models in QMRA:

- The dose–response parameters from published studies need to be evaluated in light of the observational data on which they are based; some models are based on stronger scientific evidence than others.
- The dose–response data also need to be evaluated in light of the population from which they have been collected and how representative this population is of the general population and any sensitive subpopulations within the general population.
- Although parameter uncertainty associated with the dose–response model is an important consideration within QMRA, many other forms of uncertainty (e.g. suitability of the model, representativeness of the cohort and of the pathogen) also need to be explicitly addressed.
- Low-dose approximations need to be evaluated for their suitability in a given exposure context.
- The units of dose (e.g. 50% tissue culture infective dose [$TCID_{50}$], oocysts, genome copies, focus-forming units [FFU]) are different between pathogens and need to be considered in the practical application of the models for risk assessment. In particular, were the same assay methods used for the dose–response data collection as for the exposure assessment? If not, what assumptions were made?

7.1.1 Variability in infectivity

The infectivity of a pathogen depends on its ability to pass host defences, find a site for colonization and cause an infection within the human host. Variability in infectivity can exist as a result of:

- variability between pathogen strains in their virulence and human infectivity; and
- variability between hosts in the strength of their immune response. Susceptible portions of the population include children, the elderly and the immunocompromised, who are more easily infected. In addition, short- or long-term immunity following an infection may protect portions of the population from future enteric infections.

7.1.2 Dose–response models

Various models exist for the dose–response relationship for infection and the dose–response relationship for illness when infected (the conditional relation for illness given infection). The most commonly applied models within QMRA are based on the single-hit theory: where every ingested pathogen particle is assumed to act independently and has an individual probability of causing infection (Haas, 1983; Teunis & Havelaar, 2000). These models consist of two components:

1) **Part 1.** Given the mean concentration in the exposure medium (i.e. dose), estimate the number of microorganisms ingested (a discrete number, i.e. 0, 1, 2, etc.). Usually a Poisson distribution is assumed (meaning that the microorganisms are well dispersed).
2) **Part 2.** Given the discrete number of microorganisms ingested, estimate the probability of infection. Each microorganism may be assumed to have the same probability (*r*) of causing infection (e.g. exponential dose–response model), or the probability of infection may be assumed to vary between microorganisms or hosts (e.g. if variability in infectivity is assumed to be beta distributed, the overall single-hit dose–response model becomes the Beta-Poisson model).

A useful property of the single-hit model is that there is an upper bound referred to as the maximum risk curve (Teunis & Havelaar, 2000). This curve describes the probability of infection when every ingested organism is assumed to be successful in causing an infection (exponential model with $r = 1$). In the absence of pathogen-specific infectivity information or to estimate the worst case for a particularly virulent pathogen strain, the maximum risk curve can be used as an upper bound.

When the dose is calculated as the discrete number of pathogens ingested (rather than the mean dose in the exposure medium), it is appropriate to use a conditional dose–response model (Haas, 2002) rather than the frequently used exponential or Beta-Poisson model, as only Part 2 (above) is needed.

Alternative approaches to dose–response modelling exist and are being increasingly proposed. For example, Messner, Berger & Nappier (2014) proposed a fractional dose–response model for describing norovirus infectivity in which an individual's probability of infection (Part 2 above) is assumed to be either exactly 0 or exactly 1. Based on the existing observational data for norovirus, and applying the Akaike information criterion (AIC), the fractional Poisson model was preferred over the Beta-Poisson model. Such alternative model structures may prove successful in describing the infectivity and immune response behaviour associated with different pathogens.

In the implementation of dose–response models within QMRA, it is often desirable to apply a low-dose approximation in order to simplify the calculations. In the low-dose region, the relationships are close to linear, and in many cases a linear approximation is reasonable. In the application of dose–response models in the drinking-water guidelines (GDWQ Table 7.4; WHO, in preparation), a low-dose simplification to the single-hit model has been adopted. Only the probability of infection from a single organism is considered; the probability of exposure to more than one organism per day is assumed to be so low that it is ignored in the dose–response relationship. At very low concentrations, this simplification is valid; however, it is important that at higher pathogen concentrations (mean dose approximately 0.3 microorganisms), the full single-hit (exponential or Beta-Poisson) model should be used to account for the risk of infection associated with exposure to more than one organism.

7.1.3 Published dose–response relationships

Many studies have been undertaken in which dose–response models have been fitted to experimental data. Published parameter values from these studies have been widely applied within QMRA in a diverse range of contexts. Each of these models has different sources of uncertainty, as they are based on different data (different doses, different measures of dose, different hosts and different sample sizes), and modelling studies have applied a range of different assumptions to produce an acceptable fit. An appreciation of these details is valuable for selecting a dose–response model and subsequently interpreting the representativeness of the resulting risk calculations. Further details of the dose–response relationships for the following pathogens are provided in Annex D:

- *Campylobacter jejuni*
- *E. coli* O157:H7
- Enteroviruses
  - Echovirus-12
  - Coxsackievirus B4

- Adenovirus
- Rotavirus
- Norovirus
- *Giardia lamblia*
- *Cryptosporidium parvum* and *C. hominis*

7.2 Disability-adjusted life years

Whereas all of the enteric pathogens considered in the context of this document may lead to gastroenteric symptoms, the duration and severity of illness (and likelihood of long-term sequelae) vary between pathogens. For managing water safety, it is desirable to give priority to pathogens that lead to the greatest burden of disease. The DALY is a summary measure of population health that incorporates the different severities and durations associated with different illnesses. The DALY has been applied as a metric within the WHO guidelines in order to provide a different relative weight to pathogens based on severity of disease outcomes (Havelaar & Melse, 2003).

One DALY represents the loss of one healthy life year. For each health outcome, DALYs are calculated as the sum of the years lost due to premature mortality and the years of productive life lost due to disability for incident cases of ill-health conditions. The DALY is an indicator of the time lived with a disability and the time lost due to premature mortality:

$$\text{DALY} = \text{YLL} + \text{YLD} \quad \text{Eq. 7.1}$$

$$= \text{Years of life lost} + \text{Years living with a disability}$$

For each hazard or health risk, the following is needed:

- identification of disease outcomes to be considered (construct the outcome tree);
- number of cases for each outcome in the population (estimate the probability associated with each outcome);
- duration of response (years of life lost); and
- severity of response (1 for death).

DALY weightings for a selection of reference pathogens developed for the guidelines are summarized in Table 7.1. Detailed examples demonstrating the calculation of DALYs for several pathogens in the Netherlands are given in Kemmeren et al. (2006). Labite et al. (2010) applied QMRA to evaluate the health impact associated with a range of potential interventions for an urban water system in Accra, Ghana. They applied the DALY metric for five reference pathogens for the local context. The assumptions and calculation of disease burden per case are summarized in Table 7.2.

Table 7.1 Summary of disease burden estimates for different drinking-water contaminants

Pathogen	Disease burden per 1 000 cases		
	YLD	YLL	DALY
Cryptosporidium parvum	1.34	0.13	1.47
Campylobacter spp.	3.2	1.4	4.6
Shiga toxin–producing *E. coli* O157	13.8	40.9	54.7
Rotavirus			
High-income countries	2.0	12	14
Low-income countries	2.2	480	482
Hepatitis A virus			
High-income countries, 15–49 years	5	250	255
Low-income countries	3	74	77

Source: Havelaar & Melse (2003)

7.3 Secondary transmission and immunity

Secondary transmission and immunity are more advanced QMRA inputs that give the most comprehensive look at disease burdens on the population from waterborne exposure. Secondary transmission refers to secondary infections that occur in the population due to human-to-human contact with persons infected by a waterborne pathogen. Immunity refers to the ability of humans to not become ill following exposure to a disease agent if they have been exposed to that agent previously. These factors have been demonstrated in QMRAs of sewage sludge risks, but remain relatively limited (Eisenberg et al., 1996, 2002, 2004). Secondary transmission and immunity are not considered within WHO risk-based guidelines for setting health-based targets and have not yet been demonstrated in QMRA for water safety planning (Medema et al., 2006; Medema & Smeets, 2009; Smeets et al., 2010).

Immunity is an important consideration for excreta-borne and waterborne pathogens for which there is now widespread use of immunization with vaccines to prevent infection and illness. Such vaccines are now widely used for polioviruses, rotaviruses and, in developed countries, hepatitis A virus. The use of such vaccines results in some members of the exposed population in some settings being more resistant to infection or protected completely from infection. Such vaccine-derived immunity in the population should be considered in selecting reference or target pathogens.

Table 7.2 Severity, duration and disease burden per case for pathogens included in a study on an urban water system in Accra, Ghana

Pathogen	Outcomes	Severity*	Duration*[a]	Disease burden per case in DALYs[b,c]
Campylobacter	**Gastroenteritis**			
	No GP (94% of all cases)	0.067	3.48 days (0.009 years)	$0.94 \times 0.067 \times 0.009 = 6.0 \times 10^{-4}$
	GP only (6%)	0.39	9.72 days (0.026 years)	$0.06 \times 0.39 \times 0.026 = 6.2 \times 10^{-4}$
	Hospitalization (9%)	0.39	14.39 days (0.039 years)	$0.09 \times 0.39 \times 0.039 = 1.3 \times 10^{-3}$
	Fatal (0.1%)	1	56 years	$0.001 \times 1 \times 56 = 0.056$
	Rea (7.1% of all cases)			
	No GP (85.8%)	0.127	222 days (0.6 years)	$0.071 \times 0.858 \times 0.127 \times 0.6 = 4.6 \times 10^{-3}$
	GP only (12%)	0.21	222 days (0.6 years)	$0.071 \times 0.12 \times 0.21 \times 0.6 = 1.07 \times 10^{-3}$
	Hospitalization (2.20%)	0.37	222 days (0.6 years)	$0.071 \times 0.022 \times 0.37 \times 0.6 = 3.47 \times 10^{-4}$
	Total			0.0621
Salmonella	**Gastroenteritis**			
	No GP (94% of all cases)	0.067	5.58 (0.015)	$0.94 \times 0.067 \times 0.015 = 9.6 \times 10^{-4}$
	GP only (6%)	0.393	10.65 (0.029)	$0.06 \times 0.393 \times 0.029 = 6.8 \times 10^{-4}$
	Hospitalization (9%)	0.393	16.15 (0.044)	$0.09 \times 0.393 \times 0.044 = 1.5 \times 10^{-3}$
	Fatal (0.1%)	1	56	$0.001 \times 1 \times 56 = 0.056$
	Rea (8%)			
	No GP (85.8%)	0.127	222 days (0.6 years)	$0.08 \times 0.858 \times 0.127 \times 0.6 = 5.2 \times 10^{-3}$
	GP only (12%)	0.21	222 days (0.6 years)	$0.08 \times 0.12 \times 0.21 \times 0.6 = 1.2 \times 10^{-3}$
	Hospitalization (2.20%)	0.37	222 days (0.6 years)	$0.08 \times 0.022 \times 0.37 \times 0.6 = 3.9 \times 10^{-4}$
	Total			0.0628
Rotavirus	Mild diarrhoea (85.6% of all cases)	0.10	7 days (0.02 years)	$1 \times 0.856 \times 0.10 \times 0.02 = 0.002$
	Severe diarrhoea (14.4% of all cases)	0.23	7 days (0.02 years)	$1 \times 0.144 \times 0.23 \times 0.02 = 6.6 \times 10^{-4}$
	Death	1	56 years	1 × 0.007 (death) × 56 = 0.392
	Total			0.39

Table 7.2 Severity, duration and disease burden per case for pathogens included in a study on an urban water system in Accra, Ghana (continued)

Pathogen	Outcomes	Severity*	Duration*[a]	Disease burden per case in DALYs[b,c]
Cryptosporidium	Watery diarrhoea	0.067	7.2 days (0.02 years)	$1 \times 0.067 \times 0.02 = 1.3 \times 10^{-3}$
	Death	1	22.5 years	1 × 0.0041 (death) × 22.5 = 0.09
	Total			0.09
Ascaris	Intestinal obstruction, population	0.024	35 days (0.1 years)	$1 \times 0.024 \times 0.1 = 2.4 \times 10^{-3}$
	Contemporaneous cognitive deficit (5% of all cases)	0.006	28 days (0.08 years)	$1 \times 0.05 \times 0.006 \times 0.08 = 2.4 \times 10^{-5}$
	Death	1	56	1 × 0.0008 (death) × 56 = 0.045
	Total			0.05

GP: general practitioner; Rea: reactive arthritis

* The severity weights and duration of disease, rotavirus and *Cryptosporidium* were from Havelaar & Melse (2003). Severity weights and duration of the outcome due to *Salmonella* and *Campylobacter* infections were from Kemmeren et al. (2006). Duration of disease following *Ascaris* infection was taken from Bundy et al. (1997), while the severity weight was taken from Lopez et al. (2006).

[a] The years of life lost following death from *Campylobacter*, rotavirus and *Ascaris* was taken to be the life expectancy at birth of Ghana – death at age of 1 year (57 – 1 = 56); as for *Cryptosporidium*, this was calculated from the human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) group, who already have a reduced life expectancy.

[b] The source of case fatalities included: *Salmonella* and *Campylobacter* (Haas, Rose & Gerba, 1999), rotavirus (Howard, Pedley & Tibatemwa, 2006), *Cryptosporidium* (based on the 10% HIV/AIDS prevalence in Accra), *Ascaris* (Crompton, 1999).

[c] DALYs = Number (of symptomatic cases) × severity weight × duration in years.

Source: adapted from Labite et al. (2010)

In risk characterization, the information from the exposure assessment and the health effects assessment is combined to generate a quantitative measure of risk.

The calculations performed during risk characterization need to be driven by the purpose and scope of the assessment, as defined during problem formulation (Chapter 5). Simply combining all the information into a single distribution of risk (encompassing a range of sources of variability and uncertainty) may not be informative. Alternatively, the careful modelling of scenarios can be extremely valuable for teasing out risk drivers and identifying opportunities for risk management. Scenarios can also facilitate the uncertainty and variability analyses and the sensitivity analysis to be undertaken during risk characterization (discussed below in section 8.3).

When combining the information from the exposure and health effects assessments, it is essential to consider consistency. Are the exposure and health effects assessments focused on the same hazards, the same population or population subgroups and the same time frame? Given (from Chapters 6 and 7) the strong reliance within QMRA on the measurement of indicators and surrogates and the application of dose–response data drawn from healthy adults or outbreak situations, consistency issues can be easily overlooked. In addition, as highlighted in section 7.1, it is important to ensure that the units of exposure are consistent with those used in the dose–response relationship.

8.1 Quantitative measures of risk

Quantitative measures of risk must combine, in some form, an expression of the two quantitative components of risk – namely, some measure of the probability of risk occurring; and the size of the impact should that risk occur. Quantitative measures of risk can be either deterministic (meaning single values or point estimates) or probabilistic (meaning that probability distributions are used to describe model variables and outcomes). Depending on the purpose of the assessment, different statistics may be applied for quantifying risk. Some considerations related to different statistics for describing risk are given in Table 8.1.

The risk can be defined by a range of outcomes, including infection, illness and DALYs (see Chapter 7), and may be defined for different time scales, including a single exposure or exposure over a year. Health targets are frequently defined in terms of an annual risk, which may seem unusual, given that each exposure event in QMRA is assumed to be independent, and hence microbial risks are not cumulative (unlike the assumption for chemical risks). The annual measure of risk has two clear advantages over a single exposure or daily risk measure. Firstly, when exposures are less frequent (e.g. a few times per year as opposed to daily exposure), a higher probability of infection can be tolerated per event. Secondly, an annual target acknowledges the variability in risk and allows for a higher risk to be tolerated on some occasions (e.g. due to seasonal or event-driven factors) as long as the overall risk for the year is below the health target. Nonetheless, if the management goal is to contain potential outbreaks, a daily health target could be advantageous (Signor & Ashbolt, 2009).

Table 8.1 Statistics for quantifying risk for QMRA

Context of the QMRA	Statistics for describing risk		Comments
Evaluate the long-term average risk associated with a particular pathway	Arithmetic mean	The average of all values or the "expected value"	The mean is weighted by all outcomes, including extremely high or low values that may lead to peak risk conditions. The influence of short-term fluctuations is averaged over the long term, which has advantages and disadvantages for the interpretation of risk.
Evaluate the most likely level of risk	Mode	The most likely value to occur	Rather than be weighted by peak risk conditions, the mode and median (not technically the most likely, but rather the middle value) give an indication of the central location of the risk distribution.
	Median	50th percentile; equal number of outcomes below and above	
Conservative: e.g. Given a variable inflow water quality, what level of treatment is required?	95th percentile	The value that will be exceeded by 5% of outcomes	The upper 95th percentile can be a useful conservative point estimate. It is less valuable for selecting the most appropriate risk management options.
Describe the distribution of risk (range of values with associated probability of occurrence)	PDF	Typically a parametric distribution fitted to available data (e.g. lognormal, gamma, beta)	Rather than rely on a point estimate to describe a single characteristic of the distribution, the full distribution of risk can be informative to identify the shape of the distribution, including how high extreme values are and how frequently they occur.
Scenario analysis: evaluate how high the risk could be for a particular pathway (worst case)	Maximum (concentration) Minimum (barrier reduction)	Boundary value of the variable (maximum or minimum possible) or selected by expert opinion based on review of available data	The maximum risk conditions are useful for exploring worst case within a scenario analysis.

8.2 Probability of infection: multiple exposure events

Dose–response models typically estimate the probability of infection (or, in some cases, illness) associated with a single-exposure event. In order to consider multiple events over a longer time frame, it is necessary to combine the individual probability using the following equation:

$$P_{inf/combined} = 1 - (1 - P_{inf/single})^{N} \qquad \text{Eq. 8.1}$$

where $P_{inf/combined}$ is the probability of one or more infections over N exposure events and $P_{inf/single}$ is the single-event probability of infection. When considering the annualized probability of infection associated with drinking-water exposure, N is typically set equal to 365 to represent daily exposure.

When combining infection probabilities associated with different event conditions, the following equation is used:

$$P_{inf/combined} = 1 - \prod_{i=1}^{m}(1 - P_{inf/i})^{N_i} \qquad \text{Eq. 8.2}$$

where $P_{inf/i}$ is the probability of infection associated with event i (of a total of m events to be considered in the analysis), which occurs N_i times over the period for which the combined risk $P_{inf/combined}$ is calculated.

An example of combining infection probabilities associated with different event conditions is given in Box 8.1 and section B4 in Annex B.

Box 8.1 Identification of the factors that drive risk: example of reliability

Evidence suggests that many improved drinking-water supplies in low-income regions suffer from poor reliability (Wang & Hunter, 2010). Hunter, Zmirou-Navier & Hartemann (2009) investigated what impact poor reliability may have on achieving health improvement targets, using a QMRA to assess the potential impact of interruptions in water supplies that forced people to revert to drinking raw (untreated) water.

Relying on data from the literature (building on the case-study reported by Howard, Pedley & Tibatemwa [2006]), models were constructed for three waterborne pathogens common in Africa: rotavirus, *Cryptosporidium* and enterotoxigenic *E. coli*. Infection risk was calculated for consumption of raw water (probability of infection on bad days when treatment is interrupted, $P_{inf/badd}$) and treated water (probability of infection when treatment is operational, $P_{inf/day}$) separately. To assess the impact of failure in supply for *N* days over the year, the annual risk of one or more infections per year ($P_{inf/year}$) was calculated using the equation:

$$P_{inf/year} = 1 - (1 - P_{inf/day})^{365-N} (1 - P_{inf/badd})^{N} \qquad \text{Eq. 8.3}$$

Selected results are given in Tables 8.2 and 8.3. Risk of infection by the target pathogens was substantially greater on days when people revert to raw water consumption. Over the course of a few days of raw water consumption, the annual health benefits attributed to consumption of water from an improved supply will be almost lost.

Table 8.2 Daily probability of infection when consuming treated water ($P_{inf/day}$) and when consuming raw water ($P_{inf/badd}$) following a supply failure

	Daily infection probability		
	E. coli	*Cryptosporidium*	Rotavirus
$P_{inf/day}$	1.90×10^{-6}	0.002 6	0.005 6
$P_{inf/badd}$	0.13	0.46	0.86

Table 8.3 Annual risk of infection for three pathogens for consuming only treated water and having to consume raw water for various numbers of days because of supply failure

Number of days in the year reliant on raw water because of supply failure	Annual probability of infection from drinking-water		
	E. coli	*Cryptosporidium*	Rotavirus
0	0.001	0.611	0.872
1	0.128	0.791	0.982
2	0.240	0.888	0.997
3	0.337	0.940	1.000
4	0.422	0.968	1.000
5	0.496	0.983	1.000
6	0.560	0.991	1.000
7	0.616	0.995	1.000
8	0.665	0.997	1.000
9	0.708	0.999	1.000
10	0.745	0.999	1.000

This study demonstrated (with very limited local data) the importance of reliability and provided a strong basis for arguing that agencies responsible for implementing improved drinking-water provisions will not make meaningful contributions to public health targets if those systems are subject to poor reliability.

Source: Hunter et al. (2009)

When calculating the mean or mode probability of infection associated with multiple exposure events, it is suitable to use the mean or mode P_{inf} as an input to equations 8.1 and 8.2. When the individual event probability of infection is not a constant value, percentiles of risk cannot be directly applied with equation 8.1 to calculate the same percentile associated with the combined risk. For example, assume that the upper or 95th percentile of the combined risk ($P_{inf/combined}$) was of interest. Applying the 95th percentile of $P_{inf/single}$ in equation 8.1 would assume that the 95th percentile occurred for every single-exposure event; this would overestimate the 95th percentile of the combined risk. To overcome this, the combined probability of infection can be quantified using a random sampling approach. A single random sample of the combined probability is estimated using the equation:

$$P_{inf/combined} = 1 - \prod_{j=1}^{N}(1 - \text{Random}[P_{inf/single}]) \qquad \text{Eq. 8.4}$$

where Random[$P_{inf/single}$] is a random sample from the distribution of $P_{inf/single}$. A sample from the full distribution of the $P_{inf/combined}$ can be estimated by repeating the simulation many thousands of times, such as by Monte Carlo simulation. The upper (95th) quantile of the random sample will then represent the 95th percentile of the $P_{inf/combined}$.

8.3 Evaluate the impact of uncertainty and variability

The concepts of uncertainty and variability were introduced in Chapter 4. The key principle within risk characterization is to evaluate the impact of uncertainty and variability in the input parameters on the overall risk estimate – in particular, how they may influence the recommendations for risk management that ensue from the assessment.

An important step is to ensure that the relevant sources of uncertainty and variability are identified, including the extent to which each factor is quantitatively accounted for in the calculations. For those factors that have not been incorporated quantitatively, it is important to review if and how they may be expected to influence the result. In reporting a result that incorporates both variability and uncertainty, it is important to consider if:

- the distribution is describing the variability in risk over time (e.g. regarding a distribution of annual risk, the 95th percentile would be exceeded once every 20 years; or regarding daily risk, the 95th percentile would be exceeded on approximately 18 days per year);
- the distribution is describing the variability in risk between individuals (e.g. five in every 100 people would be exposed to a risk level higher than the 95th percentile); and
- the distribution is describing uncertainty in risk (e.g. the confidence level that the annual risk is lower than the 95th percentile is 95%).

Even though their nature is different, it can be practical to characterize uncertainty and variability in the same process, provided that they are identified clearly, because of their different implications for risk management.

8.3.1 Uncertainty analysis

There is no universal method for uncertainty analysis. Annex B provides an example of how to incorporate variability and uncertainty in deterministic and probabilistic QMRA. It addresses risks associated with a water supply system that relies on surface water prone to occurrence of *Cryptosporidium.* Parts A–C of Annex B describe a deterministic assessment of a linear QMRA model. This approach lends itself best for the evaluation of the impact of individual parameters, and less for the evaluation of multiple, interrelated parameters. Consider, for instance, the possibility of a peak event in source water leading to high *Cryptosporidium* concentrations, coinciding with highly turbid water that leads to a low recovery efficiency of the *Cryptosporidium* detection method and a reduced efficiency of the coagulation/filtration process. To address such combinations, a plausible upper-bound or worst-case analysis is possible, combining the upper or extreme bounds of each input parameter into the risk model. This is a relatively simple approach, but it also reduces all available information about variability and uncertainty to a point estimate, leads to very high risk estimates and may imply costly risk management options. Therefore, although such approaches are useful to explore the upper boundaries of the risk scenarios, worst-case analysis is less valuable for selecting the most appropriate risk management options. Probabilistic methods are preferred, as they incorporate most of the available information and knowledge and allow the assessment of the relative contributions of uncertainty and variability to the risk estimate. Annex B (Parts D–F) shows how to apply probabilistic methods to account for variability and uncertainty by fitting probability distributions to the input parameter data and how to combine these distributions in the risk characterization with methods such as Monte Carlo analysis (Annex B; Haas, Rose & Gerba, 1999; Haas & Eisenberg, 2001) or second-order Monte Carlo analysis, in which variability and uncertainty can be separated (Nauta, 2000).

Assumptions are needed where knowledge is too limited or even absent. It is therefore difficult to objectively assess the impact of assumptions on the outcome of the QMRA. In any case, the assumptions and the rationale for the choices that were made should be documented. When possible, their impact on the overall outcome of the QMRA should be explored, particularly on the selection of the alternative risk management options. Ignoring assumptions in the uncertainty analysis can lead to false confidence. Methods are emerging in climate science (van der Sluis, 2005) and environmental health science (Kloprogge, van der Sluijs & Petersen, 2011) to systematically analyse the priority and value of assumptions and their impact on the risk assessment results, using expert and stakeholder evaluation.

8.3.2 Sensitivity analysis

Sensitivity analysis is recommended for (Frey, Mokhtari & Zheng, 2004):

- prioritization of potential control points in the system;
- identification of key sources of uncertainty and variability;
- refinement and verification of the QMRA model; and
- conditional analysis of the QMRA model ("what if" scenario analysis and identification of factors contributing to high exposure or risk).

Although the objectives of sensitivity analysis and uncertainty analysis are different, the results of these two types of analysis are usually not independent; an input parameter with a high degree of uncertainty may be highly influential in the sensitivity analysis. Methods for undertaking sensitivity analysis have been reviewed by Frey & Patil (2002) and Frey, Mokhtari & Zheng (2004). There are mathematical, statistical and graphical methods available. The selection of the appropriate method for sensitivity analysis depends on a range of factors, such as the objective, level of detail required, detail and type of data and information used in the model and available software/resources. The basic method for sensitivity analysis is to change the input value for individual model components and determine the magnitude of change of the estimated risk.

For deterministic models and for a crude sensitivity screening in probabilistic models, the nominal range sensitivity analysis (NRSA) is appropriate in many cases. In NRSA, the value of one model input parameter is varied over its range of plausible values while all other model inputs are kept at their baseline value, and the magnitude of change of the calculated model output is recorded. The larger the change in the output, the more sensitive the model is to the input parameter. Repeating this for each input parameter generates an overview of the relative sensitivities of the model to each of the inputs. NRSA works best with linear models. This approach is not suitable to handle variations in multiple parameters when there may be interactions between model inputs. Statistical methods (analysis of variance, correlation analysis) can identify the effect of simultaneous interactions among multiple input parameters. Distributions for model inputs can be propagated through a model using a variety of sampling techniques, such as Monte Carlo simulation, Latin hypercube sampling and other sampling methods (Cullen & Frey, 1999).

Table 8.4 shows an example of NRSA of a QMRA model developed for the risk of infection of *Cryptosporidium* from drinking-water produced from a surface water source (as used in Annex B), with the baseline and plausible upper bounds for each of the input parameters. The sensitivity value shows the magnitude of change in the risk of infection. It is clear that in this example, the model is most sensitive to peak concentrations of *Cryptosporidium* in source water and the periods of reduced treatment performance. Hence, risk mitigation options focused on source water protection and improved reliability of treatment.

Table 8.4 Sensitivity analysis of a QMRA of a surface water supply (see also Annex B)

Stage/barrier	Sensitivity value tested	Type of variability/ uncertainty	Baseline value	Plausible upper bound	Sensitivity value
Source water	Highest *Cryptosporidium* concentrations ($n \cdot L^{-1}$)	Variability	0.1	8	80
	Lowest observed recovery efficiency of detection method	Parameter uncertainty	0.41	0.25	1.64
	Large fraction of environmental oocysts is not infectious to humans	Assumption uncertainty	1	0.3	0.3
Barrier A (coagulation + filtration)	Lowest observed decimal reduction	Variability	0.003 2 (= 2.5 log)	0.079 (= 1.1 log)	25
	Bacillus monitoring has missed more sporadic (1/year) severe barrier failures	Assumption uncertainty	1	10[a]	1.02
Barrier B (chlorination)	Not considered, as chlorination is not effective against *Cryptosporidium*	–	–	–	–
Consumption of unheated tap water	High consumption (L)	Variability	0.15	1	6.7
Dose–response	Maximum infections per organism	Assumption uncertainty	0.4	1	2.5

n: number

[a] 1 of 365 days.

Box 8.2 illustrates the application of sensitivity analysis for identification of the most important sources of variability and uncertainty within a model for evaluating the risk of viral infection associated with consumption of wastewater-irrigated lettuce.

Box 8.2 Assessment of risk of exposure to enteric viruses via lettuce crops

Sensitivity analysis is a powerful tool for assessing and developing strategies for the management of risks. In most practical situations, it is not enough to simply characterize what the risk is; rather, an understanding of the most important data gaps and drivers of risk is needed. Petterson (2002) investigated the risks associated with enteric viruses present in irrigation water, which attach to the lettuce during spray irrigation and are able to persist on the lettuce crop to consumption. The components of the exposure assessment are illustrated in Fig. 8.1. The conceptual pathway was favoured for two reasons. Firstly, data on the occurrence of viruses on lettuce crops were unavailable; however, enumeration data were available on the concentration of enteric viruses in secondary treated effluent. Secondly, the model would allow for the risk drivers, risk management options (including irrigation practices and crop withholding periods) and the most important sources of uncertainty in the model to be identified.

Box 8.2 Assessment of risk of exposure to enteric viruses via lettuce crops (continued)

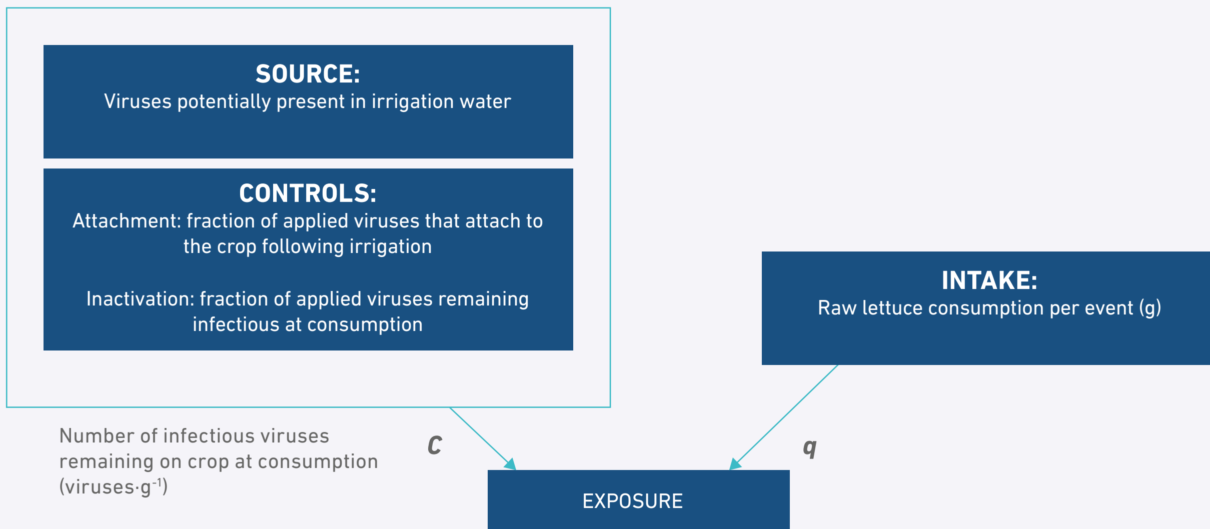


Fig. 8.1 Schematic of exposure pathway applied for the assessment of enteric virus risk via consumption of wastewater-irrigated lettuce crops (adapted from Petterson & Ashbolt, 2015)

The exposure was quantified using the equation:

$$\text{Exposure} = c \times v \times f \times S(t) \times q \qquad \text{Eq. 8.5}$$

where:

Exposure is the mean dose per exposure event

c is the concentration of viruses in the irrigation water applied to the crop (virus·L^{-1})

v is the volume of irrigation water applied to the crop (L·g^{-1})

f is the fraction of applied viruses that attach to the crop following irrigation

$S(t)$ represents the fraction of viruses remaining infectious at time t following irrigation

q is the quantity of crops consumed per event.

The quantification of model components was undertaken based on scientific data; a summary is given in Table 8.5. A best estimate and an extreme (reasonable worst-case) estimate for each of the model parameters were made. Two measures of sensitivity were applied for the exposure assessment:

1) **Step characteristic** (SC) is applied to quantify the main determinants of risk. The SC_k indicates the $\log_{10}$ reduction or increase in the number of organisms at step k relative to the previous step in the model and is given by: $SC_k = \log\left(\frac{N_k}{N_{k-1}}\right)$ where N_k is the number of organisms per unit mass/volume at step k

2) **Factor sensitivity** (FS) is applied to identify the importance of uncertainty and/or variability at each step k in the model and is given by: $FS_k = \log\left(\frac{N_k\text{extreme}}{N_k\text{average}}\right)$

Box 8.2 Assessment of risk of exposure to enteric viruses via lettuce crops (continued)

Table 8.5 Best and extreme estimates of model parameters applied in the exposure assessment of virus exposure via wastewater-irrigated lettuce crops

Model component	"Best" estimate	"Extreme" estimate
Virus occurrence: c	2.6 viruses·L^{-1}[a]	470 000 viruses·L^{-1}[b]
Virus attachment[c]: f	0.024	0.071
Virus inactivation[d]: $S(t)$	h_1 = 2.5 d^{-1}	h_1 = 2.0 d^{-1}
Biphasic inactivation	h_2 = 0.5 d^{-1}	h_2 = 0.3 d^{-1}
$S(t) = a \cdot h_1 + (1-a)\ h_2$	a = 0.12%	a = 0.96%
Consumption per event[e]: q	100 g	300 g

[a] Californian data set of enteroviruses by cell culture used by Asano et al. (1992); Petterson & Ashbolt (2001); Hamilton et al. (2006).

[b] Upper limit reviewed by Yates & Gerba (1998).

[c] Maximum likelihood (best) and upper 95% credible interval (extreme) based on modelling undertaken by Petterson (2002).

[d] Maximum likelihood (best) and conservative 95% credible interval (extreme) based on modelling undertaken by Petterson, Teunis & Ashbolt (2001).

[e] Reference values applied by Petterson (2002).

The SC and FS were calculated for each component of the process (the volume of irrigation water applied was kept constant at 1 mL·g^{-1}, and consumption was assumed to occur at 14 days following last irrigation). The results are included in Table 8.6. The SC indicates the importance of each component in the model for driving the exposure. For the current model assumptions, virus inactivation was the most important barrier for reducing exposure to infectious viruses, identifying the importance of climatic conditions and crop withholding periods for controlling risk.

Table 8.6 Step characteristic and factor sensitivity calculated for the lettuce crop exposure model

Process step	SC	FS
Virus occurrence: c	–	5.49
Virus attachment: f	−1.6	0.45
Virus inactivation: $S(t)$	−6.2	2.20
Consumption per event: q	–	0.48

The FS quantifies the influence of variability and/or uncertainty associated with each component of the model. Virus concentration in the irrigation water is uncertain and subject to wide fluctuations and therefore was the most important driver of uncertainty in the model outcomes (FS = 5.49). The second most sensitive factor was the quantification of virus survival (FS = 2.20), whereas the uncertainty in virus attachment (FS = 0.45) and crop consumption (FS = 0.48) had a comparably smaller effect on the uncertainty in the calculated exposure.

As an additional measure of sensitivity, the critical limits were determined for each of the exposure variables. The critical limit was defined as the value at which the probability of infection exceeds 1×10^{-6} per consumption event (based on an annual risk of infection of 1×10^{-4} and an exposure duration of 100 days per year), with all other components kept at their best estimates. The results are given in Table 8.7. The risk was below the critical limit when all variables were set equal to their best estimates. The results show the magnitude of change in each variable that could be tolerated. When all other values were set equal to their best estimates, consumption of lettuce would need to exceed 16 kg in order for the tolerable risk limit to be exceeded.

Box 8.2 Assessment of risk of exposure to enteric viruses via lettuce crops (continued)

Table 8.7 Critical limits for tested variables

Tested variable	Critical value[a]
Virus occurrence (inactivation parameters equal to best estimate)	425 viruses·L^{-1}
Virus occurrence (inactivation parameters equal to conservative 90th percentiles)	3 viruses·L^{-1}
Probability of attachment (*f*)	Maximum $f = 1$ risk is still tolerable
Subpopulation size (*a*)	17%
Time since final irrigation (inactivation parameters equal to best estimate)	4.3 days
Time since final irrigation (inactivation parameters set equal to conservative 90th percentiles)	14 days
Consumption per event (*q*)	16 kg

[a] Value at which probability of infection exceeds 1×10^{-6} per consumption event (based on annual risk of infection of 1×10^{-4}), with all other components kept at their best estimates unless otherwise noted.

FS was also calculated for the dose–response model by comparing each candidate model against the maximum risk curve (exponential model with $r = 1$). The results are shown in Table 8.8. For infectious viruses such as rotavirus and adenovirus, application of the maximum risk curve (rather than a virus-specific model) would have a relatively small impact (0.22 and 0.27) in comparison with other factors in the exposure assessment. For less infectious enteric viruses, such as echovirus, application of the maximum risk curve would overestimate the risk by nearly 3 $\log_{10}$.

Table 8.8 Factor sensitivity associated with selected dose–response model for quantifying risk

Dose–response model	FS[a]
Rotavirus	0.22
Echovirus	2.75
Adenovirus	0.27

[a] FS of calculated infection risk from maximum risk curve (exponential model with $r = 1$).

The simple deterministic measures of sensitivity calculated in this example are very informative for understanding the drivers in the model in terms of the relative importance of the different barriers (SC) and the most important sources of variability and uncertainty (FS). If additional monitoring or data gathering were to be undertaken to inform the characterization of exposure, improved understanding of the magnitude of the virus concentration in irrigation waters would be much more beneficial than improved understanding of virus attachment behaviour or consumer consumption patterns.

Source: Petterson (2002)

9 HOW CAN QMRA SUPPORT WATER SAFETY MANAGEMENT?

In this chapter, the ways in which QMRA can support water safety management both at the broad regulatory level and for site-specific system management are described, with reference to examples from the peer-reviewed published literature.

The specific role of QMRA to support water safety management, particularly in the drinking-water context, has been demonstrated through many and various examples presented in the literature (Medema & Smeets, 2009; Smeets et al., 2010; Petterson & Ashbolt, 2016). This chapter provides an overview of this work, with examples expanded to include recreational and wastewater reuse pathways. The application of QMRA for the development of national regulations or guidelines is first considered, followed by four ways to support the site-specific WSP or SSP: 1) to support system understanding (“know your system”); 2) to evaluate system or pathway safety (“system assessment”); 3) to identify control points and define or evaluate monitoring requirements, including critical limits; and 4) to support system planning and incremental improvement.

9.1 Regulation and health-based treatment targets

In the GDWQ (WHO, in preparation) and the GWEG (WHO, 2006a,b), QMRA is recommended as the framework for translating a quantitative tolerable health outcome target, such as 1×10^{-6} DALY pppy, to performance targets, such as required pathogen removal by water treatment or other control measures (the recommended approach is illustrated in Fig. 9.1). The approach is founded on scientific evidence and provides a transparent process for identifying the magnitude of treatment and operational control required to achieve an identified health outcome target. The QMRA method also provides a framework that is analogous to the risk-based approach to setting guideline values for chemicals (NRC, 2009; WHO, 2010). There are, of course, uncertainties in the translation of the scientific evidence into the constrained framework, which have been discussed in this document. The Dutch Drinking Water Act (discussed in Box 9.1) requires that risk assessment be undertaken for waterborne pathogens at every water supply to demonstrate microbiologically safe water. An example of implementation of the translation of health outcome targets to operational targets in national guidelines is illustrated in Case-study 5 in Annex A, with the Australian Guidelines for Water Recycling (NWQMS, 2006).

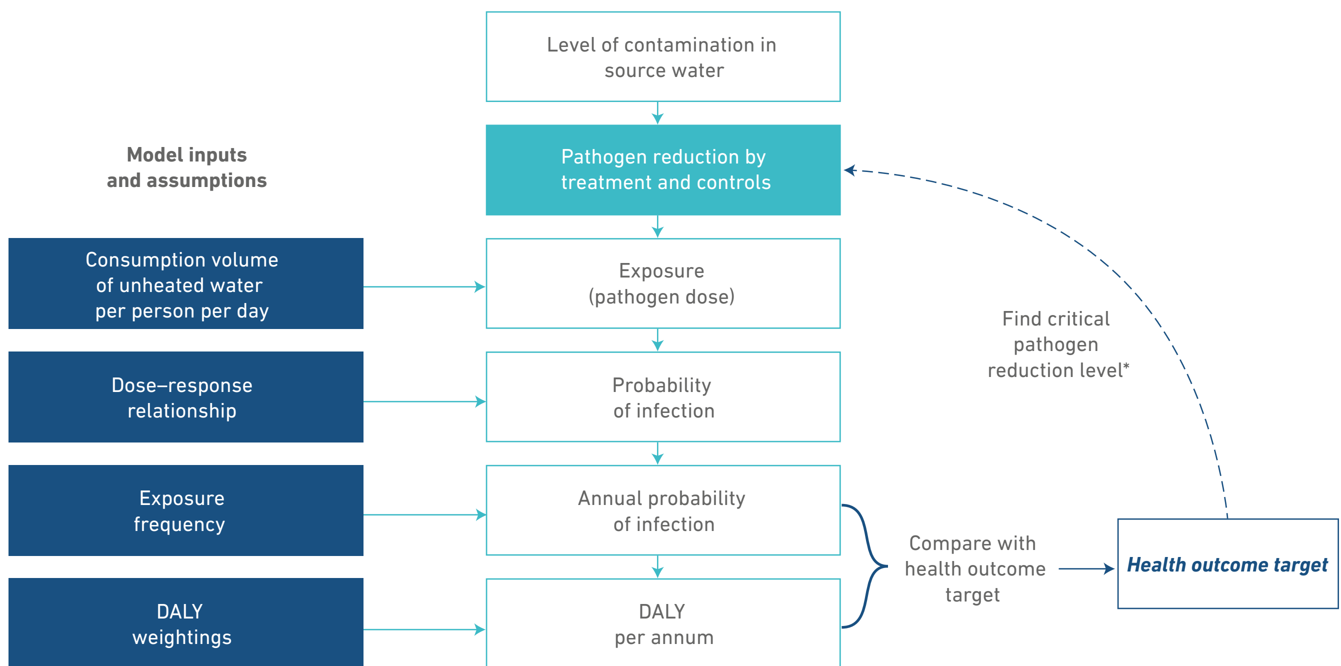


Fig. 9.1 QMRA approach for defining a performance target for treatment to meet health outcome target (adapted from Petterson, Roser & Deere, 2015)

Box 9.1 Role of regulations and application of QMRA in the Netherlands

The Dutch Drinking Water Act of 2001 set a health-based regulatory target of less than one infection per 10 000 persons per year and stipulates that risk assessment be undertaken for waterborne pathogens to demonstrate microbiologically safe water (Anonymous, 2001). If the assessed infectious risk exceeds this target value, the drinking-water company must consult with the national environmental inspector about necessary mitigation measures. In the Dutch Drinking Water Act itself, no specific directives were given on how to perform the risk assessment. This led to ambiguity regarding which data should be used and assumptions applied in the QMRA and which statistic (e.g. mean, 95th percentile) of the risk should meet the regulatory target. Therefore, in 2005, the Inspectorate Guideline 5318 was created to define requirements of the QMRA and create a "level playing field" for the water utilities (Anonymous, 2005). The introduction of this guideline prescribed how uncertainty and variability should be dealt with for quantifying risk in relation to the regulation. The guideline includes the following:

- The risk assessment should be undertaken for the reference pathogens *Campylobacter*, enteroviruses, *Cryptosporidium* and *Giardia*.
- Raw water supplies should be monitored for the reference pathogens (frequency depending on size of production) over 3 years.
- Appropriate surrogate organisms or process models for the assessment of treatment efficacy are recommended.
- Data on consumption of unheated tap water for the Netherlands are prescribed.
- The dose–response models to be used in the assessment are specified.

In addition, while the mean infection risk must be below 1 in 10 000, the guideline recommends that at least 95% of the time, the infection risk should be below the target. To support the legislative requirements, a software tool (QMRAspot) was developed by the National Institute for Public Health and the Environment (RIVM) to facilitate QMRA calculations using the methods and assumptions recommended in the Inspectorate Guideline 5318 (Schijven et al., 2011). This guideline and its associated tools have helped water utilities in the Netherlands to identify the main hazards and the important barriers to control these. It has prompted the necessity for rigorous monitoring of these critical control points and for preplanned remedial actions in the case of barrier failure. The development of the Inspectorate Guideline 5318, and how it clarified the interpretation of variability and uncertainty within the regulatory QMRA, was a critical component of the successful implementation of the Dutch Drinking Water Act of 2001.

9.2 Site-specific water and sanitation safety planning

WSPs and SSPs are the recommended frameworks for managing risks associated with a specific water supply or water reuse application. QMRA can provide valuable quantitative input into the steps of the WSP, as illustrated in Fig. 9.2. Examples of these different applications are summarized in the following sections.

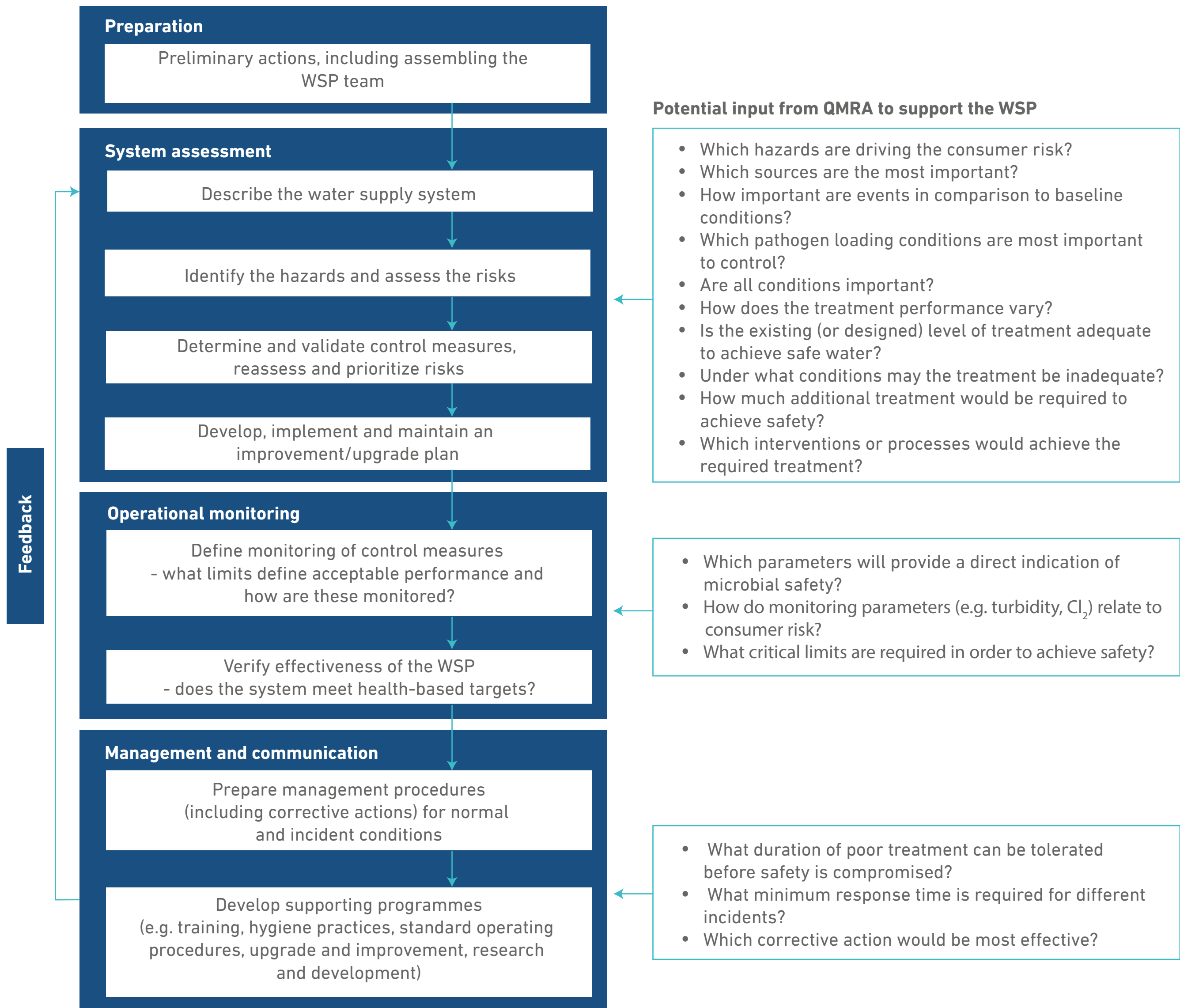


Fig. 9.2 Potential input from QMRA to support water safety planning (adapted from Petterson & Ashbolt, 2016)

9.2.1 Know your system

The underlying principle of water and sanitation safety planning is to "know your system", so as to identify the priority hazards and associated hazardous events and interventions. It can help answer questions such as: Which drinking-water supply scenario is safest? Is water a significant transmission route for gastrointestinal illness? Characterizing a system from source to exposure, gathering scientific data, selecting quantitative values (first with conservative screening-level values) and predicting the implications for public health are of great value for improving the water manager's (and WSP or SSP team's) understanding of their system. When prioritizing systems or hazards, assumptions should be openly tested against scientific data to aid in understanding the system. Many examples from wastewater reuse, recreational water and drinking-water have been published that demonstrate how the systematic approach of QMRA informed an understanding of the water system and its vulnerabilities, including:

- identification of the most hazardous risk pathways for a particular scheme (Westrell et al., 2004; Oesterholt et al., 2007; Diallo et al., 2008);
- identification of the factors that drive risk (Signor & Ashbolt, 2006; Hunter, Zmirou-Navier & Hartemann, 2009);
- identification of the most important barriers (Medema et al., 2003; Hamilton et al., 2006; Åström et al., 2007);
- identification of the most sensitive components within the risk model (Petterson & Ashbolt, 2001a; Hamilton et al., 2006; Jaidi et al., 2009);
- determining the relative significance of major, well-controlled and minor, less well-controlled routes of exposure and the moments of reduced treatment performance (Gale, 2002);
- prioritization of hazards by comparing possible pathogen and disinfection by-product risks with alternative water systems for a coastal community in the USA (Schoen et al., 2014);
- evaluation of the relative importance of different faecal sources on the overall risk to the community (Case-study 1 in Annex A: Schoen & Ashbolt, 2010, 2011; Soller et al. , 2010b, 2015); and
- demonstration of the importance of compliance on overall consumer risk (Hunter, Zmirou-Navier & Hartemann, 2009; Enger et al., 2013; Petterson, 2015).

9.2.2 System assessment

QMRA provides a clear and transparent approach for comparing system risks with a health outcome target, making it possible to evaluate whether a system or pathway is safe. This can be conducted in the format of a deterministic, screening-level risk assessment. The example in section 2.3 illustrates a deterministic, screening-level QMRA to assess the safety of a surface water supply system following the approach in the WHO GDWQ (WHO, in preparation).

Case-study 3 (Annex A) demonstrates the broad application of this approach to a large number of drinking-water supplies in France. Conservative screening-level risk assessments such as these can be very useful in identifying low-risk and high-risk scenarios. Cases have been reported where risks were evaluated to be so low in comparison with targets that the system was concluded to be acceptable (Höglund, Stenström & Ashbolt, 2002; Medema et al., 2003; van den Akker et al., 2011; Pintar et al., 2012); and, conversely, where risks were so high in comparison with targets that the system was concluded to be unacceptable and additional controls were identified as necessary (Medema et al., 2003; Steyn, Jagals & Genthe, 2004; van Heerden et al., 2005; Oesterholt et al., 2007; Schönning et al., 2007; Hunter et al., 2011; Seidu et al., 2015). However, as the risk approaches the region of the regulatory target, the decision-making process becomes more complex, and the costs and benefits of acting conservatively (e.g. constructing or including additional barriers) or undertaking further analysis must be weighed. This is illustrated in more detail in section 4.3 and the detailed quantitative example in Fig. 4.2.

Closely linked with the system assessment is the identification of the magnitude of additional controls required to achieve the defined level of safety. In the Plurinational State of Bolivia, Symonds et al. (2014) applied QMRA to assess the safety, with respect to enteric viruses, of reusing wastewater effluent from stabilization ponds. When evaluating the safety of using wastewater for irrigation, they concluded that 1–2 $\log_{10}$ additional treatment would be needed following the three-pond systems, and 2.5–4.5 $\log_{10}$ additional treatment would be needed for the upflow anaerobic sludge blanket pond system. If children were to be allowed to play in irrigated fields, then 3–5 $\log_{10}$ additional treatment would be needed. The example from Yang et al. (2015) given in Box 9.2 demonstrates the application of this principle for evaluating what controls are necessary to maintain safety following a distribution system failure.

Box 9.2 QMRA for system assessment and maintaining safety following events: examples of microbial risk modelling for mains breaks

This QMRA is a summary of Yang et al. (2015), who conducted a QMRA to derive operational performance targets for cleaning the water distribution network after repair works in the USA setting. Pathogen data were collected from the literature and pilot studies were conducted to evaluate the efficacy of flushing and disinfection methods to remove or inactivate pathogens.

Problem formulation

Given the reports on health risks associated with (repair works of) drinking-water distribution networks, the water utility wanted to develop evidence-based operational targets for cleaning the network after repair works on drinking-water distribution networks following mains breaks. A QMRA was designed to determine the significance of cleaning measures and to derive operational targets for adequate cleaning. Raw sewage was considered to be the riskiest contamination source, and a meta-analysis was done on pathogen concentration data in sewage from the scientific literature. Norovirus, *E. coli* O157 and *Cryptosporidium* were selected as reference pathogens on the basis of their incidence in the human population and in sewage and their infectivity and health effects. The health target that was used to derive the operational targets was the average risk of infection of 10^{-4} pppy.

Exposure assessment

Pathogen data from different studies were combined in a meta-analysis. The reported pathogen concentrations in sewage were variable, and hence statistical (lognormal) distributions were fitted to the collected data. The second step of the exposure assessment was to determine the amount of material that enters the distribution network during mains breaks and repair. Evaluation of the break and repair process showed that, in practice, sewage would be diluted by the potable water, and the dilution would be 99–99.99%. To evaluate the effect of cleaning the network by flushing, pilot studies were conducted to determine the removal of sand, which was selected as a conservative material, because higher flushing velocities are needed to fluidize sand particles compared with other mineral soil materials. The pilot studies produced the relationship between flushing velocity and log removal of the sand particles. Bench-scale disinfection studies were done to determine the efficacy of chlor(am)ination on the inactivation of indicator organisms (MS2 as model virus, *E. coli* as model bacterium and *Bacillus* spores as model parasite). The disinfection experiments were conducted in water matrices such as encountered in practice. For the consumption of unheated water, survey data were used, and a statistical (lognormal) distribution was fitted to the data.

Health effects assessment

Published dose–response relationships for norovirus, *Cryptosporidium* and *E. coli* O157 were used.

Risk characterization

To incorporate the variability and uncertainty in pathogen concentrations in the source, the dilution factor in the mains, the efficacy of flushing and disinfection and the dose–response parameters, a probabilistic QMRA was conducted. A Monte Carlo simulation that generated 10 000 samples of each of the model inputs to create 10 000 estimates of the risk of infection was undertaken (Fig. 9.3). A sensitivity analysis was conducted to evaluate the sensitivity of the calculated risk to the different input parameters (pathogen concentration, dilution factor, efficacy of flushing and disinfection, water consumption and dose–response parameters).

Box 9.2 QMRA for system assessment and maintaining safety following events: examples of microbial risk modelling for mains breaks (continued)

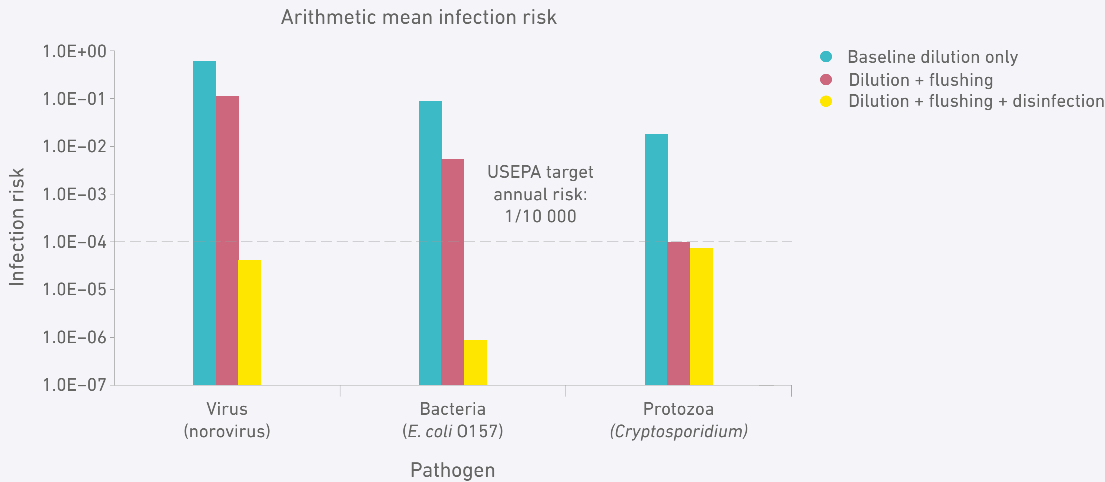


Fig. 9.3 Risk of infection for the first customer after a main break event after dilution, flushing and disinfection (adapted from Kirmeyer et al., 2014)

For each of the pathogens, the risk of infection was clearly above the health outcome target of 10^{-4} if no cleaning methods were applied. For *Cryptosporidium,* flushing at a minimum velocity of 3 $ft \cdot s^{-1}$ (0.9 $m \cdot s^{-1}$) was sufficient to reduce the risk below 10^{-4}. For norovirus and *E. coli* O157, this was not the case, and additional disinfection was needed to further reduce their concentrations. The disinfection experiments indicated that a free chlorine concentration × contact time (ct) value of 100 $mg \cdot min \cdot L^{-1}$ was needed (because of the presence of particulates) to achieve sufficient inactivation. The utility used this QMRA to translate the 10^{-4} risk of infection target to operational performance targets of flushing at >3 $ft \cdot s^{-1}$ (>0.9 $m \cdot s^{-1}$) and free chlorine disinfection at >100 $mg \cdot min \cdot L^{-1}$ to produce safe drinking-water after mains breaks.

Source: Yang et al. (2015)

9.2.3 Operational targets and setting critical limits

QMRA can provide a scientific basis for setting operational targets and critical limits for water treatment and system management. QMRA therefore allows for efficient implementation of control measures so that performance is adequate but not excessive, so as to balance the allocation of resources and costs against the required health benefit. This is of particular importance when balancing risks, such as of dosing of chemical disinfectants that lead to the formation of potentially toxic disinfection by-products, against the benefits of disinfection of drinking-water.

In particular, when a treatment target (i.e. required $\log_{10}$ reduction) has been defined by a regulation or guideline (e.g. using the approach illustrated in Fig. 9.1), it is necessary to design, monitor and control the operation of any given control measure so that the required target is achieved. Traditional end-of-pipe testing that relies only on the absence of faecal indicators is inadequate for verifying the operational integrity of the treatment train, as mentioned in Chapter 1. Monitoring of control measures therefore predominantly relies on process indicators (e.g. chlorine residual) rather than microbial sampling where operational targets and critical limits are generally established. How well these process indicators can be correlated with treatment performance (in terms of $\log_{10}$ reduction) and hence consumer risk will be critical for further application in this area.

The concept and value of using QMRA to define operational and critical limits have been described in the literature (Westrell et al., 2004; Nilsson et al., 2007; Smeets et al., 2010), although few examples exist. In particular, defining the chemical dose and/or residence time required to achieve the necessary $\log_{10}$ reduction means that suboptimal performance can be easily identified through measurable operational criteria, and corrective action can be subsequently instigated. Box 9.3 gives an example of how QMRA can be used to derive critical limits for chlorination. In another example, Åström et al. (2007) used QMRA to evaluate the adequacy of an *E. coli*–driven operational limit (in collaboration with upstream event notification) on opening/closing a raw water intake at a drinking-water treatment plant in Sweden.

Box 9.3 Setting critical limits

A treatment system can be designed to provide exactly the right level of treatment to meet the performance targets. However, in practice, the risk manager needs to account for variations and inaccuracies in order to run a practical and stable process. A simulated example illustrates how QMRA can be used to set critical limits and setpoints for a water disinfection process. The performance target for the disinfection process was 3 log_{10} inactivation of viruses. Chlorine concentration, contact time (flow), temperature and pH were monitored on-line and translated into virus log_{10} inactivation levels using an evidence-based disinfection model of the relationship between these conditions (including the residence time distribution in the contact chambers) and the log_{10} inactivation of enteroviruses. The required chlorine residual for 3 log_{10} inactivation was 1.7 $mg \cdot L^{-1}$, as determined by this disinfection model. This level was considered the critical limit: if the process was constantly run at exactly this concentration, the performance target would be met; conversely, if there was an exceedance of the critical limit, the target would not be met. In practice, the chlorine dose would vary as a result of operational variation. Chlorine dosing was controlled by an automatic control loop. The dosing rate was adjusted if the measured chlorine concentration deviated from the setpoint. As a result, the chlorine residual level would typically vary between the operational limits if the system was working within specifications, as indicated in Fig. 9.4. At hour 20, the chlorine dosing pump became clogged, resulting in a chlorine level below the operational limit. This triggered an alarm, and the operator was able to clean the pump and restore normal operation. This event did not affect the average treatment performance in such a way that the performance target would not be met, as the critical limit was not exceeded. At hour 80, the chlorine dosing pump failed again; this time, however, the operator was not in time to restore the system before the critical limit was exceeded. The short period of time during which disinfection was ineffective reduced the average treatment efficacy, as the critical limit was exceeded. Therefore, the operator needed to take corrective actions, such as starting emergency chlorination of the distributed water, in order to restore achievement of the performance target.

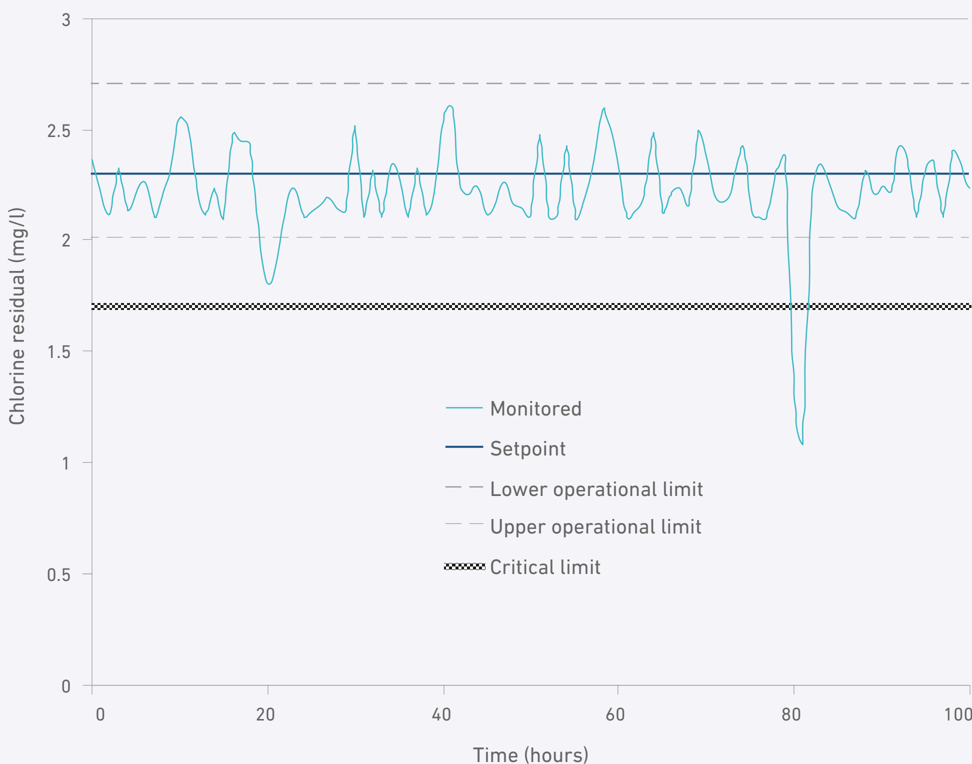


Fig. 9.4 Simulated example of critical limits, operational limits, setpoints and monitoring of a chlorine disinfection process (Smeets et al., 2010)

9.2.4 System planning and development

QMRA can be a valuable planning tool for assessing the next best steps for incremental improvement in health outcomes. Some examples include the following:

- Howard, Pedley & Tibatemwa (2006; see Box 9.4) used a deterministic QMRA for the assessment of different sources of drinking-water and subsequent prioritization of infrastructure investment to maximize health outcomes.
- In Case-study 2 (Annex A), Weir et al. (2011) used QMRA to evaluate the expected risk reduction associated with different interventions at a water spray park.
- In Ghana, QMRA was applied to estimate the risk to human health from various sources of drinking-water and evaluate the cost-effectiveness of different interventions (Machdar et al., 2013).
- Various water and sanitation pathways were assessed to compare the impact of different interventions in Ghana (Labite et al., 2010) and Uganda (Katukiza et al., 2014).
- Microbial risks were included within a sustainability framework for assessing urban water system options in Australia (Lundie et al., 2008; Kobayashi et al., 2015) and in the USA (Schoen et al., 2014).
- When investigating a high incidence of *Cryptosporidium* in a south-western Ontario community in Canada, Pintar et al. (2012) applied QMRA to effectively rule out the drinking-water pathway as the cause, providing an effective communication tool and prioritizing future work.
- QMRA was applied to select among household water treatment devices in the developing context (Petterson, 2015).
- The cost–benefit analysis undertaken by the United States Environmental Protection Agency (USEPA) is summarized in Case-study 4 (Annex A).

Box 9.4 QMRA to determine the health risk associated with different types of water supplies

This QMRA is summarized from Howard, Pedley & Tibatemwa (2006), who conducted a QMRA on the water supply of Kampala, Uganda. The source water is Lake Victoria, and the water is treated by conventional treatment and chlorination and transported via a piped network to homes (20% of the population) or standpipes (52% of the population). Twenty-eight per cent of the population was collecting water from local springs. Data on thermotolerant coliforms were available for each of these types of water supply (at the tap and in household containers).

Problem formulation

The utility wanted to know the health risk associated with the three different types of water supply in Kampala. Based on the disease burden in Uganda and human and animal faecal contamination of the source water, *E. coli* O157 was selected as the reference pathogen for enteric bacteria. The end-point of the QMRA was DALYs, and local disease data were used to derive a local disease burden in DALYs per case.

Exposure assessment

No pathogen data were available. Data on faecal indicators from lake water, from the distribution network taps in homes and standpipes, from the springs and from household containers were available and used to estimate pathogenic *E. coli* concentrations in spring water, household containers, taps, standpipes and lake water and to estimate removal of the microorganisms through the water treatment processes (Table 9.1). Data on *E. coli* were used for *E. coli* O157 (using literature data on the fraction [8%] of pathogenic *E. coli* among total *E. coli* in water). Consumption of water was assumed to be 1 L per day.

Health effects assessment

Published dose–response relationships were used. The fraction of the population served by the different types of water supply was known and taken into account. The fraction was discounted for the use of water from multiple sources. The total population was considered vulnerable because immunity is short-lived. As indicated, local disease burden estimates were made based on information about the severity of illness and life expectancy in Uganda.

Box 9.4 QMRA to determine the health risk associated with different types of water supplies (continued)

Risk characterization

A deterministic QMRA was conducted, using point estimates for each of the variables. The concentration of *E. coli* was translated to the concentration of enterohaemorrhagic *E. coli*. The concentration in drinking-water was combined with the volume of drinking-water consumed (Table 9.1) to calculate the exposure of the population using each type of water supply. Using the dose–response and disease burden values (in the same way as in the WHO GDWQ Table 7.4 and the example in Chapter 2), the disease burden from enterohaemorrhagic *E. coli* O157 was calculated for the populations served by each of the water supply types (last row of Table 9.1). The risk was below the 10^{-6} DALY pppy guideline in the water leaving the treatment plant, but in the water arriving in the households, the risk was over 400-fold higher and did not meet the DALY guideline anymore. This highlighted the need for direct investment in improved operation and maintenance of the network, so this was a priority for risk management. In addition, the water from the springs and in the household containers collected at the standpipes yielded the highest estimated disease burden and therefore highlighted the need for effective household water treatment and safe storage in the interim and in the longer term to increase access to piped water in homes in order to protect public health. While another risk assessment approach (e.g. the risk matrix approach) may have resulted in the same conclusions, the QMRA provided further reassurance to justify investment priorities to maximize health outcomes.

Table 9.1 QMRA of enterohaemorrhagic *E. coli* in drinking-water from a surface water supply in Kampala, with water leaving the treatment works and water arriving in the households

	Treatment	Households	Standpipes	Spring
Raw water quality (n/L)	7.5			
Log removal conventional	3			
Log removal chlorination	2			
Drinking-water quality (n/L)	7.50×10^{-5}	0.1	2.3	10.6
Consumption of unheated drinking-water (L)	1	1	1	1
Exposure by drinking-water	7.50×10^{-5}	0.1	2.3	10.6
Dose–response (probability of infection per organism)	1.00×10^{-3}	1.00×10^{-3}	1.00×10^{-3}	1.00×10^{-3}
Risk of infection per day	7.50×10^{-8}	1.00×10^{-4}	2.30×10^{-3}	1.06×10^{-2}
Risk of infection per year	2.74×10^{-5}	3.65×10^{-2}	8.40×10^{-1}	3.87
Risk of diarrhoeal disease given infection	0.25	0.25	0.25	0.25
Risk of diarrhoeal disease	6.84×10^{-6}	9.13×10^{-3}	2.10×10^{-1}	9.67×10^{-1}
Maximum disease burden per case	0.32	0.32	0.32	0.32
Susceptible (exposed) population (%)	31	10	42	28
Disease burden	6.8×10^{-7}	2.9×10^{-4}	2.8×10^{-2}	8.7×10^{-2}

Source: adapted from Howard, Pedley & Tibatemwa (2006)

9.3 Principles of QMRA that are relevant to risk management

Many examples have been cited that show the potential value of QMRA; in practice, however, the value of the outcomes for decision support relies emphatically on the way in which uncertainty and variability have been considered and accounted for in the overall analysis. Decision-makers need to be provided with an appreciation of the representativeness of the QMRA results and the magnitude of the uncertainty surrounding any ensuing recommendations.

There are several characteristics associated with QMRA investigations that are useful for risk management. The following principles relate to each of the steps of the harmonized QMRA framework contained within this document:

- Problem formulation (Chapter 5):
  - The QMRA is driven by the risk management need – the decision (and the required level of certainty for that decision) has driven the problem formulation.
  - The QMRA is fit-for-purpose in terms of scope and level of detail.
  - The QMRA is based on exposure scenarios consistent with the risk management options, local context and data constraints.
- Exposure assessment (Chapter 6):
  - Factors influencing variability associated with each component of the exposure are considered and documented.
  - The data and information used to quantify exposure are the most appropriate for use and weighed against the information in the scientific literature.
- Health effects assessment (Chapter 7):
  - The representativeness of published dose–response models is evaluated.
  - The QMRA is focused on the population exposed. As appropriate, it will give consideration to susceptible population subgroups and life stages.
- Risk characterization (Chapter 8):
  - The quantities of risk calculated are driven by the problem and directly relate to the risk management decision.
  - The influence of uncertainties on the outcome of the assessment is evaluated.

The following overarching principles support the development of QMRA investigations that provide a reliable and defensible basis for risk management:

- **Transparency:** QMRA requires good record-keeping. All assumptions should be referenced and supported by the available data and the scientific literature. Procedures that are used for data processing should be explained and recorded. All steps, data inputs, key assumptions, limitations and decisions should be documented, along with associated rationale and justification. These should be clearly conveyed along with the results of the QMRA, and the potential impact on the results should be discussed, wherever possible. It is also good practice to construct computer models that clearly follow the structure of the conceptual model. Good documentation of the QMRA results in transparency, which enables peers, risk managers and other stakeholders to understand the basis for model development, data selection and processing, and assumptions. It is of great value in uncertainty analysis, verification, external review and communication of the model and results to the risk manager.
- **Verification:** The risk assessor should verify all elements of the QMRA to ensure that the model, data used, assumptions, calculations and uncertainties are appropriate. Where alternative assumptions could have been made, they should be evaluated together with other uncertainties (see also Chapter 8).
- **Review:** Peer review is an independent, external review of all elements in the QMRA (conceptual model, computational model, data, assumptions, uncertainty analysis and conclusions) to ensure that it meets the standards of the scientific community (Lammerding, 2007; Boone et al., 2010). It is advisable to start the peer review in the early stages of the QMRA, to facilitate proper data and model selection.

This transparency and the implementation of a robust verification and review process will help to increase confidence that the QMRA approach taken is adequate for the risk management questions that the QMRA addresses and that the conclusions about risk management options are credible.

10 | CONCLUSIONS AND NEXT STEPS

In this document, the following factors for facilitating the effective implementation of QMRA for practical system management have been identified and explored:

- There is considerable value in adopting a harmonized approach to QMRA across the water-related exposure routes: drinking-water, wastewater and excreta reuse, and recreational water. These exposure routes should not be considered in isolation, as they are linked within environmental systems, and public health risk management actions need to be balanced between these community exposure routes. In addition, scientific considerations associated with the occurrence, fate/transport, exposures and health impacts overlap between exposure routes, adding to the value of a harmonized approach. Nonetheless, it is understood that current jurisdictional issues may hamper such a broad approach.
- The value of the outcomes of any QMRA rests firmly on the appropriate interpretation of the scientific evidence and hence the representativeness of the model input assumptions. It is therefore important for risk assessors to incorporate variability and uncertainty analyses into risk assessment and communicate clearly with risk managers regarding the implications of variability and uncertainty on the risk outcomes.
- QMRA has been shown to support a wide range of water safety management decisions. The case-studies included within this document (see Annex A) highlight the value of QMRA for the evaluation of different hazards and alternative risk management options, determining risk management priorities, selecting the most effective interventions, supporting cost–benefit analysis of risk management actions and setting of health-based performance targets. These case-studies also illustrate that QMRA can support different levels of risk management – from (supra)national water safety regulations, strategies and policies to utility management priorities.
- Overall, QMRA is an aid to provide evidence-based understanding of hazards and their control in water systems and therefore provides valuable insight for risk management. Risk management usually requires delicate trade-offs: proportionate allocation of resources to the most effective risk management measures or the weighing of water safety management against other societal demands. QMRA is the tool to provide the evidence base to support proportionate, efficient water safety management. It is essential that risk assessors work together with risk managers from the problem formulation stage of the QMRA to ensure that the assessment is fit-for-purpose.

To support broader application of QMRA, additional case-study applications from countries or regions with few resources and high diarrhoeal disease burdens are needed. Although QMRA has been identified to be of value in the developing country context, particularly for prioritizing interventions and resource allocation, practical case-studies from such settings are still limited and largely theoretical. The majority of the skills and experience within QMRA have been developed in resource-rich settings with low disease burdens and cannot give due consideration to the practical realities of data-limited environments with high disease burdens.

Whereas expert opinion is essential in the selection of model inputs and assumptions, it is increasingly recognized that a more systematic approach to evaluating data sources may be valuable. Currently, such a system does not exist for environmental data and QMRA, and therefore development of a suitable system to weigh evidence associated with environmental data and QMRA would facilitate the robustness of QMRA.

The application of QMRA for water systems would be facilitated by a public, web-based compendium of:

- case-studies that demonstrate the value of QMRA in different settings and for different purposes;
- research studies that provide additional scientific basis and methodology for QMRA;
- input data, such as pathogen concentrations in different water matrices, data on pathogen removal by different barriers under different conditions, data on exposure volumes and dose–response data.

Examples of such compendia are available: QMRAwiki (http://qmrawiki.canr.msu.edu/) and for food risk management (http://www.foodrisk.org). This guidance document has assembled and discussed the three elements (QMRA science and methodology, input data and case-studies) and can serve as a starting point for such a compendium.

REFERENCES

- Abrahams D, Pennington A, Scott-Samuel A, Doyle C, Metcalfe O, den Broeder L et al. (2004). European policy health impact assessment – A guide. European Commission, Health and Consumer Protection Directorate-General (http://ec.europa.eu/health/ph_projects/2001/monitoring/fp_monitoring_2001_a6_frep_11_en.pdf, accessed 5 May 2016).
- Anonymous (1998). [This is how Netherlands eats 1998. Results of a food consumption survey 1998.] The Hague: Voedingscentrum (in Dutch).
- Anonymous (2001). [Decision of January 9, 2001 to amend the Drinking Water Decree in connection with the Directive on the quality of water intended for human consumption.] Official Journal of the Kingdom of the Netherlands. 31, 1e53 (in Dutch).
- Anonymous (2005). [Inspectorate Guideline Analysis microbiological safety of drinking water.] VROM Inspectorate. Item code 5318 (in Dutch).
- Asano T, Leong LYC, Rigby MG, Sakaji RH (1992). Evaluation of the California wastewater reclamation criteria using enteric virus monitoring data. Water Sci Technol. 26(7–8):1513–24.
- Ashbolt NJ, Reidy C, Haas CN (1997). Microbial health risk at Sydney's coastal bathing beaches. In: Proceedings of the 17th Australian Water and Wastewater Association meeting, vol. 2. Melbourne: Australian Water and Wastewater Association; 104–11.
- Ashbolt NJ, Schoen ME, Soller JA, Roser DJ (2010). Predicting pathogen risks to aid beach management: the real value of quantitative microbial risk assessment (QMRA). Water Res. 44(16):4692–703.
- Åström J, Petterson S, Bergstedt O, Pettersson TJ, Stenström TA (2007). Evaluation of the microbial risk reduction due to selective closure of the raw water intake before drinking water treatment. J Water Health. 5(Suppl 1):81–97.
- Bartram J, Corales L, Davison A, Deere, D, Drury, D, Gordon B et al. (2009). Water safety plan manual. Step-by-step risk management for drinking-water suppliers. Geneva: World Health Organization (http://apps.who.int/iris/bitstream/10665/75141/1/9789241562638_eng.pdf, accessed 9 April 2016).
- Bassett J, Nauta M, Lindqvist R, Zwietering M (2012). Tools for microbiological risk assessment. Brussels: International Life Sciences Institute.
- Berg T, Viberg U (2003). [Drinking water – A study of consumption and behaviour.] Linköping: Linköping University, Department of Management and Economics (in Swedish).
- Beuken R, Reinoso M, Sturm S, Kiefer J, Bondelind M, Åström J et al. (2008). Identification and description of hazards for water supply systems – A catalogue of today's hazards and possible future hazards, updated version. Nieuwegein: TECHNEAU (https://www.techneau.org/fileadmin/files/Publications/Publications/Deliverables/D4.1.4.pdf, accessed 26 August 2015).
- Bichai F, Smeets PWMH (2012). Using QMRA-based regulation as a water quality management tool in the water security challenge: experience from the Netherlands and Australia. Water Res. 47:7315–26.
- Blumenthal UJ, Fleisher JM, Esrey SA, Peasey A (2001). Epidemiology: a tool for the assessment of risk. In: Fewtrell L, Bartram J, editors. Water quality: guidelines, standards and health – assessment of risk and risk management for water-related infectious disease. London: IWA Publishing on behalf of the World Health Organization; 135–60.
- Boone I, Van der Stede Y, Aerts M, Mintiens K, Daube G (2010). Quantitative microbial risk assessment: methods and quality assurance. Vlaams Diergeneeskd Tijdschr. 79:367–81.

- Borchardt MA, Spencer SK, Kieke BA, Lambertini E, Loge FJ (2012). Viruses in nondisinfected drinking water from municipal wells and community incidence of acute gastrointestinal illness. Environ Health Perspect. 120(9):1272–9.
- Bundy DAP, Chan MS, Medley GF, Jamison D, de Silva NR, Savioli L (1997). Intestinal nematode infections. In: Health priorities and burden of disease analysis: methods and applications from global, national and subnational studies. Cambridge (MA): Harvard University Press.
- Codex Alimentarius Commission (1999). Principles and guidelines for the conduct of microbiological risk assessment. CAC/GL 30-1999. Adopted 1999, Amendments, 2012, 2014. Rome: Food and Agriculture Organization of the United Nations, Joint FAO/WHO Food Standards Programme, Codex Alimentarius Commission (http://www.fao.org/docrep/004/y1579e/y1579e05.htm, accessed 8 April 2016).
- Collier SA, Stockman LJ, Hicks LA, Garrison LE, Zhou FJ, Beach MJ (2012). Direct healthcare costs of selected diseases primarily or partially transmitted by water. Epidemiol Infect. 140(11):2003–13.
- Crabtree KD, Gerba CP, Rose JB, Haas CN (1997). Waterborne adenovirus – A risk assessment. Water Sci Technol. 35(11–12):1–6.
- Craig DL, Fallowfield HJ, Cromar NJ (2003). Effectiveness of guideline faecal indicator organism values in estimation of exposure risk at recreational coastal sites. Water Sci Technol. 47(3):191–8.
- Craun GF, Brunkard JM, Yoder JS, Roberts VA, Carpenter J, Wade T et al. (2010). Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. Clin Microbiol Rev. 23(3):507–28.
- Crompton DW (1999). How much helminthiasis is there in the world? J Parasitol. 85:397–403.
- Cullen A, Frey HC (1999). Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs. New York (NY): Plenum Press.
- Cullen AC, Small MJ (2004). Uncertain risk. The role and limits of quantitative assessment. In: McDaniels T, Small MJ, editors. Risk analysis and society: an interdisciplinary characterization of the field. Cambridge: Cambridge University Press.
- Dangendorf F (2003). [Water-related gastrointestinal infections in Germany? An analysis of the geographical extension of gastro-intestinal infections with regard to the structure of the drinking-water distribution system using the example of Rheinisch Bergischer Kreis] [dissertation]. Bonn: Rheinische Friedrich-Wilhelms-Universität Bonn (in German).
- Deere D, Stevens M, Davison A, Helm G, Dufour A (2001). Management strategies. In: Fewtrell L, Bartram J, editors. Water quality: guidelines, standards and health – assessment of risk and risk management for water-related infectious disease. London: IWA Publishing on behalf of the World Health Organization; 257–88.
- De Keuckelaere A, Jacxsens L, Amoah P, Medema GJ, McClure P, Jaykus LA et al. (2015). Zero risk does not exist: lessons learned from microbial risk assessment related to use of water and safety of fresh produce. Compr Rev Food Sci Food Saf. 14(4):387–410.
- Diallo MB, Anceno AJ, Tawatsupa B, Houpt ER, Wangsuphachart V, Shipin OV (2008). Infection risk assessment of diarrhea-related pathogens in a tropical canal network. Sci Total Environ. 407(1):223–32.
- Dorevitch S, Panthi S, Huang Y, Li H, Michalek AM, Pratap P et al. (2011). Water ingestion during water recreation. Water Res. 45(5):2020–8.
- Dufour AP, Evans O, Behymer TD, Cantú R (2006). Water ingestion during swimming activities in a pool: a pilot study. J Water Health. 4(4):425–30.
- DWI (1996). Tap water consumption in England and Wales: findings from the 1995 National Survey. London: Drinking Water Inspectorate.
- EHD (1981). Tap water consumption in Canada. Ottawa: Health Canada, Environmental Health Directorate, Health Protection Branch.
- Eisenberg JN, Seto EYW, Olivieri AW, Spear RC (1996). Quantifying water pathogens risk in an epidemiological framework. Risk Anal. 16(4):549–63.
- Eisenberg JN, Brookhart MA, Rice G, Brown M, Colford JM Jr (2002). Disease transmission models for public health decision making: analysis of epidemic and endemic conditions caused by waterborne pathogens. Environ Health Perspect. 110(8):783–90.
- Eisenberg JN, Soller JA, Scott J, Eisenberg DM, Colford JM Jr (2004). A dynamic model to assess microbial health risks associated with beneficial uses of biosolids. Risk Anal. 24(1):221–36.
- Enger KS, Nelson KL, Rose JB, Eisenberg JN (2013). The joint effects of efficacy and compliance: a study of household water treatment effectiveness against childhood diarrhea. Water Res. 47(3):1181–90.
- Environmental Protection Agency (2010). European Communities (Drinking Water) (No. 2). Regulations 2007. A handbook on the implementation of the regulations for water services authorities for private water supplies. Wexford: Environmental Protection Agency (https://www.epa.ie/pubs/advice/drinkingwater/privatewatersupplieshandbook/Section 10.pdf, accessed 26 August 2015).

- Eyles R, Niyogi D, Townsend C, Benwell G, Weinstein P (2003). Spatial and temporal patterns of *Campylobacter* contamination underlying public health risk in the Taieri River, New Zealand. J Environ Qual. 32(5):1820–8.
- FAO/WHO (2003). Hazard characterization for pathogens in food and water. Guidelines. Rome: Food and Agriculture Organization of the United Nations; and Geneva: World Health Organization (Microbial Risk Assessment Series, No. 3; http://apps.who.int/iris/bitstream/10665/42693/1/9241562374.pdf, accessed 4 May 2016).
- Ferguson C, Charles K, Deere D (2009). Quantification of microbial sources in drinking-water catchments. Crit Rev Environ Sci Technol. 39(1):1–40.
- Ferguson CM, Croke BF, Beatson PJ, Ashbolt NJ, Deere DA (2007). Development of a process-based model to predict pathogen budgets for the Sydney drinking water catchment. J Water Health. 5(2):187–208.
- Fewtrell L, Bartram J, editors (2001). Water quality: guidelines, standards and health. Assessment of risk and risk management for water-related infectious disease. London: IWA Publishing on behalf of the World Health Organization.
- Foekema H, Engelsma O (2001). [A different consumption pattern. Water consumption at home in 2001.] The Hague: VEWIN/NIPO (in Dutch).
- Frey HC, Patil SR (2002). Identification and review of sensitivity analysis methods. Risk Anal. 22(3):553–78.
- Frey HC, Mokhtari AS, Zheng J (2004). Recommended practice regarding selection, application and interpretation of sensitivity analysis methods applied to food safety process risk models. North Carolina State University, Raleigh, NC. Prepared for the Office of Risk Assessment and Cost–Benefit Analysis, United States Department of Agriculture, Washington, DC.
- Gale P (2002). Using risk assessment to identify future research requirements. J Am Water Works Assoc. 94(9):30–8.
- Gale P (2005). Land application of treated sewage sludge: quantifying pathogen risks from consumption of crops. J Appl Microbiol. 98(2):380–96.
- Genthe B, Rodda N (1999). Application of health risk assessment techniques to microbial monitoring data. Pretoria: Water Research Commission.
- Girones R, Ferrûs MA, Alonso JL, Rodriguez-Manzano J, Calgua B, Corrêa Ade A et al. (2010). Molecular detection of pathogens in water – the pros and cons of molecular techniques. Water Res. 44(15):4325–39.
- Gofti-Laroche L, Potelon JL, da Silva E, Zrnirou D (2001). Description of drinking water intake in French communities (E.M.I.R.A. study). Rev Epidemiol Sante Publique 49(5):411–22.
- Haas CN (1983). Estimation of risk due to low doses of microorganisms: a comparison of alternative methodologies. Am J Epidemiol. 118(4):573–82.
- Haas CN (2002). Conditional dose–response relationships for microorganisms: development and application. Risk Anal. 22(3):455–63.
- Haas CN, Eisenberg JNS (2001). Risk assessment. In: Fewtrell L, Bartram J, editors. Water quality: guidelines, standards and health. Assessment of risk and risk management for water-related infectious disease. London: IWA Publishing on behalf of the World Health Organization; 161–83.
- Haas CN, Rose JB, Gerba CP (1999). Quantitative microbial risk assessment. New York (NY): John Wiley & Sons, Inc.
- Haas CN, Rose JB, Gerba CP (2014). Quantitative microbial risk assessment, second edition. New York (NY): John Wiley & Sons, Inc.
- Havelaar AH, Melse JM (2003). Quantifying public health risk in the WHO Guidelines for Drinking-water Quality: a burden of disease approach. Bilthoven: National Institute for Public Health and the Environment (RIVM Report 734301022/2003; http://www.who.int/water_sanitation_health/dwq/rivmrep.pdf?ua=1, accessed 4 May 2016).
- Hamilton AJ, Stagnitti F, Premier R, Boland AM, Hale G (2006). Quantitative microbial risk assessment models for consumption of raw vegetables irrigated with reclaimed water. Appl Environ Microbiol. 72(5):3284–90.
- Hara-Kudo Y, Takatori K (2011). Contamination level and ingestion dose of foodborne pathogens associated with infections. Epidemiol Infect. 139(10):1505–10.
- Haring BJA, Karres JJC, Poel P, van der Zoeteman BCJ (1979). [Research into the use habits of drinking water consumption in the Netherlands.] H2O. 12(10):212–6 (in Dutch).
- Havelaar A, Melse JM (2003). Quantifying public health risk in the WHO Guidelines for Drinking-water Quality: a burden of disease approach. Bilthoven: National Institute for Public Health and the Environment (RIVM Report 734301022).
- Havelaar AH, De Hollander AEM, Teunis PFM, Evers EG, Van Kranen HJ, Versteegh JFM et al. (2000). Balancing the risks and benefits of drinking water disinfection: disability adjusted life-years on the scale. Environ Health Perspect. 108(4):315–21.
- Hijnen WAM, Beerendonk EF, Medema GJ (2006). Inactivation credit of UV radiation for viruses, bacteria and protozoan (oo) cysts in water: A review. Water Res. 40(1):3-22

- Hijnen WAM, Medema GJ (2010). Elimination of micro-organisms by drinking water treatment processes. A review, third edition. London: IWA Publishing.
- Höglund C, Stenström TA, Ashbolt N (2002). Microbial risk assessment of source-separated urine used in agriculture. Waste Manage Res. 20(2):150–61.
- Hopkin SM, Ellis JC (1980). Drinking water consumption in Great Britain. A survey of drinking habits with special reference to tap-water based beverages. Medmenham: Water Research Centre.
- Howard G, Pedley S, Tibatemwa S (2006). Quantitative microbial risk assessment to estimate health risks attributable to water supply: can the technique be applied in developing countries with limited data? J Water Health. 4(1):49–65.
- Howard G, Ahmed MF, Shamsuddin AJ, Mahmud SG, Deere D (2006). Risk assessment of arsenic mitigation options in Bangladesh. J Health Popul Nutr. 24(3):346–55.
- Hunter PR, Zmirou-Navier D, Hartemann P (2009). Estimating the impact on health of poor reliability of drinking water interventions in developing countries. Sci Total Environ. 407(8):2621–4.
- Hunter PR, Payment P, Ashbolt NJ, Bartram J (2003). Assessment of risk. In: Assessing microbial safety of drinking water. Improving approaches and methods. London: IWA Publishing on behalf of the World Health Organization and the Organisation for Economic Co-operation and Development; 79–109 (http://www.who.int/water_sanitation_health/publications/assessing-microbial-safety-of-drinking-water/en/, accessed 10 May 2016).
- Hunter PR, Hughes SG, Woodhouse S, Syed Q, Verlander NQ, Chalmers RM et al. (2004). Sporadic cryptosporidiosis case–control study with genotyping. Emerg Infect Dis. 10(7):1241–9.
- Hunter PR, de Sylor MA, Risebro HL, Nichols GL, Kay D, Hartemann P (2011). Quantitative microbial risk assessment of cryptosporidiosis and giardiasis from very small private water supplies. Risk Anal. 31(2):228–36.
- ILSI (2000). Revised framework for microbial risk assessment. An ILSI Risk Science Institute workshop report. Washington (DC): International Life Sciences Institute, ILSI Press.
- Jaidi K, Barbeau B, Carriere A, Desjardins R, Prevost M (2009). Including operational data in QMRA model: development and impact of model inputs. J Water Health. 7(1):77–95.
- Katukiza A, Ronteltap M, van der Steen P, Foppen J, Lens P (2014). Quantification of microbial risks to human health caused by waterborne viruses and bacteria in an urban slum. J Appl Microbiol. 116(2):447–63.
- Kemmeren JM, Mangen M-JJ, van Duynhoven YTHP, Havelaar AH (2006). Priority setting of foodborne pathogens, disease burden and costs of selected enteric pathogens. Bilthoven: National Institute for Public Health and the Environment (RIVM Report 330080001/2006; http://www.rivm.nl/dsresource?objectid=rivmp:13228&type=org&disposition=inline&ns_nc=1, accessed 9 April 2016).
- Kirmeyer GJ, Thomure TM, Rahman R, Marie JL, LeChevallier MW, Yang J et al. (2014). Effective microbial control strategies for main breaks and depressurization. Denver (CO): Water Research Foundation.
- Kloprogge P, van der Sluijs J, Petersen AC (2011). A method for the analysis of assumptions in model-based environmental assessments. Environ Model Softw. 26:289–301.
- Kobayashi Y, Peters GM, Ashbolt NJ, Shiels S, Khan SJ (2015). Assessing burden of disease as disability adjusted life years in life cycle assessment. Sci Total Environ. 530–531:120–8.
- Krumpel E, Nelson K (2013). Comparing microbial water quality in an intermittent and continuous piped water supply. Water Res. 47(14):5176–88.
- Labite H, Lunani I, van der Steen P, Vairavamoorthy K, Drechsel P, Lens P (2010). Quantitative microbial risk analysis to evaluate health effects of interventions in the urban water system of Accra, Ghana. J Water Health. 8(3):417–30.
- Lammerding AM (2007). Using microbiological risk assessment (MRA) in food safety management: summary report of a workshop held in October 2005 in Prague, Czech Republic. Brussels: ILSI Europe.
- LeChevallier M, Au KK (2004). Water treatment and pathogen control: process efficiency in achieving safe drinking water. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/water-treatment-and-pathogen-control/en/, accessed 9 April 2016).
- Lloyd B, Bartram JK (1991). Surveillance solutions to microbiological problems in water quality control in developing countries. Water Sci Technol. 24(2):61–75.
- Lopez AD, Mathers CD, Ezzati M, Jamison TD, Murray CJL (2006). Global burden of disease and risk factors. New York (NY): Oxford University Press.
- Lundie S, Ashbolt N, Peters G, Livingston D, Lai E, Karman E et al. (2008). Sustainability framework. Part A: Methodology for evaluating the overall sustainability of urban water systems. Part B: A review comparing the WSAA sustainability framework to the Gold Coast Waterfuture process. Melbourne: Water Services Association of Australia (WSAA Occasional Paper No. 17).
- Machdar E, van der Steen NP, Raschid-Sally L, Lens PNL (2013). Application of quantitative microbial risk assessment to analyze the public health risk from poor drinking water quality in a low income area in Accra, Ghana. Sci Total Environ. 449:134–42.

- MacKenzie WR, Hoxie NJ, Proctor ME, Gradus MS, Blair KA, Peterson DE et al. (1994). A massive outbreak in Milwaukee of *Cryptosporidium* infection via the public water supply. N Engl J Med. 331:161–7.
- Masago Y, Oguma K, Katayama H, Hirata T, Ohgaki S (2004). *Cryptosporidium* monitoring system at a water treatment plant, based on waterborne risk assessment. Water Sci Technol. 50(1):293–9.
- Medema G, Smeets P (2009). Quantitative risk assessment in the water safety plan: case studies from drinking water practise. Water Sci Technol Water Supply. 9(2):127–32.
- Medema GJ, Ketelaars HAM, Hoogenboezem W, Rijs GBJ, Schijven JF (2001). *Cryptosporidium* and *Giardia*: occurrence in sewage, manure and surface water. Amsterdam: Association of River Waterworks (RIWA); 171.
- Medema GJ, Hoogenboezem W, van der Veer AJ, Ketelaars HAM, Hijnen WAM, Nobel PJ (2003). Quantitative risk assessment of *Cryptosporidium* in surface water treatment systems. Water Sci Technol. 47(3):241–7.
- Medema G, Loret J-C, Stenström TA, Ashbolt N (2006). Quantitative microbial risk assessment in the water safety plan. Final report on the EU MicroRisk Project. Brussels: European Commission.
- Medema GJ, Teunis PF, Blokker M, Deere D, Davidson D, Charles P et al. (2009). Risk assessment of *Cryptosporidium* in drinking water. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/cryptoRA/en/, accessed 9 April 2016).
- Mena KD, Gerba CP, Haas CN, Rose JB (2003). Risk assessment of waterborne coxsackievirus. J Am Water Works Assoc. 95(7):122–31.
- Messner MJ, Berger P, Nappier SP (2014). Fractional Poisson – A simple dose–response model for human norovirus. Risk Anal. 34(10):1820–9.
- Ministry of Health (2014). Water safety plan guides for drinking water supplies. Wellington: New Zealand Ministry of Health (http://www.health.govt.nz/publication/water-safety-plan-guides-drinking-water-supplies, accessed 26 August 2015).
- Mons MN, van der Wielen JM, Blokker EJ, Sinclair MI, Hulshof KF, Dangendorf F et al. (2007). Estimation of the consumption of cold tap water for microbiological risk assessment: an overview of studies and statistical analysis of data. J Water Health. 5(Suppl 1):151–70.
- Morgan MG, Henrion M (1990). Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge: Cambridge University Press.
- Muirhead RW, Collins RP, Bremer PJ (2006a). Numbers and transported state of *Escherichia coli* in runoff direct from fresh cowpats under simulated rainfall. Lett Appl Microbiol. 42(2):83–7.
- Muirhead RW, Collins RP, Bremer PJ (2006b). Interaction of *Escherichia coli* and soil particles in runoff. Appl Environ Microbiol. 72(5):3406–11.
- Nauta M (2000). Separation of uncertainty and variability in quantitative microbial risk assessment models. Int J Food Microbiol. 57:9–18.
- Nilsson P, Roser D, Thorwaldsdotter R, Petterson S, Davies C, Signor R et al. (2007). SCADA data and the quantification of hazardous events for QMRA. J Water Health. 5(Suppl 1):99–105.
- NRC (1983). Risk assessment in the federal government: managing the process. Washington (DC): National Research Council.
- NRC (2009). Science and decisions: advancing risk assessment. Washington (DC): National Research Council, National Academies Press.
- NWQMS (2006). National Water Quality Management Strategy – National guidelines for water recycling: managing health and environmental risks. Report of the National Resource Management Ministerial Council, Environment Protection and Heritage Council and Australian Health Ministers' Conference (http://www.environment.gov.au/system/files/resources/044e7a7e-558a-4abf-b985-2e831d8f36d1/files/water-recycling-guidelines-health-environmental-21.pdf, accessed 9 April 2016).
- Oesterholt F, Martijnse G, Medema G, van der Kooij D (2007). Health risk assessment of non-potable domestic water supplies in the Netherlands. J Water Supply Res Technol – Aqua. 56(3):171–9.
- Ohno K, Asami M, Matsui Y (2013). Questionnaire survey on water ingestion rates for various types of liquid and the seasonal differences between summer and winter. In: Proceedings of the Annual Meeting of the Society for Risk Analysis, Baltimore, Maryland, USA, 8–11 December. McLean (VA): Society for Risk Analysis.
- Ottoson J, Stenström TA (2003). Faecal contamination of greywater and associated microbial risks. Water Res. 37(3):645–55.
- Payment P (2013). Assessing waterborne health risks through quantitative risk assessment models. Laval: Centre INRS-Institut Armand-Frappier (http://www.cwn-rce.ca/assets/Uploads/Payment-Assessing-Waterborn-Health-Risks-QMRA-2014-Report.pdf, accessed 26 August 2015).
- Petterson SR (2002). Quantitative microbial risk assessment for evaluating the viral risk from consumption of wastewater irrigated salad crops [PhD dissertation]. Sydney: University of New South Wales, School of Civil and Environmental Engineering.

- Petterson SR (2015). Application of a QMRA framework to inform selection of drinking water interventions in the developing context. Risk Anal. 36(2):203–14.
- Petterson SR, Ashbolt NJ (2001). Viral risks associated with wastewater reuse: modelling virus persistence on wastewater irrigated salad crops. Water Sci Technol. 43(12):23–6.
- Petterson SR, Ashbolt NJ (2015). Exposure assessment. In: Yates MV, Nakatsu CH, Miller RV, Pillai SD, editors. Manual of environmental microbiology, fourth edition. Washington (DC): ASM Press; 3.5.2-1–3.5.2-18. doi:10.1128/9781555818821.ch3.5.2.
- Petterson SR, Ashbolt NJ (2016). QMRA and water safety management: review of application in drinking water systems. J Water Health. doi:10.2166/wh.2016.262.
- Petterson SR, Stenström TA (2015). Quantification of pathogen inactivation efficacy by free chlorine disinfection of drinking water for QMRA. J Water Health. 13(3):625–44.
- Petterson S, Roser D, Deere D (2015). Characterizing the concentration of *Cryptosporidium* in Australian surface waters for setting health-based targets for drinking water treatment. J Water Health. 13(3):879–96.
- Petterson SR, Teunis PFM, Ashbolt N (2001). Modeling virus inactivation on salad crops using microbial count data. Risk Anal. 21(6):1097–1107.
- Pintar KD, Fazil A, Pollari F, Waltner-Toews D, Charron DF, McEwen SA et al. (2012). Considering the risk of infection by *Cryptosporidium* via consumption of municipally treated drinking water from a surface water source in southwestern Ontario community. Risk Anal. 32(7):1122–38.
- Quintero-Betancourt W, Peele ER, Rose J (2002). *Cryptosporidium parvum* and *Cyclospora cayetanensis*: a review of laboratory methods for detection of these waterborne parasites. J Microbiol Methods. 49:209–24.
- Regli S, Rose JB, Haas CN, Gerba CP (1991). Modeling the risk from *Giardia* and viruses in drinking water. J Am Water Works Assoc. 83(11):76–84.
- Risebro HL, Doria MF, Andersson Y, Medema G, Osborn K, Schlosser O et al. (2007). Fault tree analysis of the causes of waterborne outbreaks. J Water Health. 5(Suppl 1):1–18.
- Robertson B, Sinclair M, Forbes A (2000). The effect of an introductory letter on participation rates using telephone recruitment. Aust N Z J Public Health. 24(5):552.
- Robertson B, Forbes A, Sinclair M, Black J (2000). How well does a telephone questionnaire measure drinking water intake? Aust N Z J Public Health. 24(6):619–22.
- Rose JB, Haas CN, Regli S (1991). Risk assessment and control of waterborne giardiasis. Am J Public Health. 81:709–13.
- Roseberry AM, Burmaster DE (1992). Lognormal distributions for water intake by children and adults. Risk Anal. 12(1):99–104.
- Schijven JF, Teunis PF, Rutjes SA, Bouwknegt M, de Roda Husman AM (2011). QMRAspot: a tool for quantitative microbial risk assessment from surface water to potable water. Water Res. 45(17):5564–76.
- Schoen ME, Ashbolt NJ (2010). Assessing pathogen risk to swimmers at non-sewage impacted recreational beaches. Environ Sci Technol. 44(7):2286–91.
- Schoen ME, Ashbolt NJ (2011). An in-premise model for *Legionella* exposure during showering events. Water Res. 45(18):5826–36.
- Schoen ME, Xue X, Hawkins TR, Ashbolt NJ (2014). Comparative human health risk analysis of coastal community water and waste service options. Environ Sci Technol. 48(16):9728–36.
- Schönning C, Westrell T, Stenström TA, Arnbjerg-Nielsen K, Hasling AB, Hansen L et al. (2007). Microbial risk assessment of local handling and reuse of human faeces. J Water Health. 5(1):117–28.
- Scottish Executive (2006). Private water supplies: technical manual. Edinburgh: Scottish Executive.
- Seidu R, Heistad A, Amoah P, Drechsel P, Jenssen PD, Stenström TA (2008). Quantification of the health risk associated with wastewater reuse in Accra, Ghana: a contribution toward local guidelines. J Water Health. 6(4):461–71.
- Seidu R, Abubakari A, Dennis IA, Heistad A, Stenström TA, Larbi JA et al. (2015). A probabilistic assessment of the contribution of wastewater-irrigated lettuce to *Escherichia coli* O157:H7 infection risk and disease burden in Kumasi, Ghana. J Water Health. 13(1):217–29.
- Signor RS, Ashbolt NJ (2006). Pathogen monitoring offers questionable protection against drinking-water risks: a QMRA (quantitative microbial risk analysis) approach to assess management strategies. Water Sci Technol. 54(3):261–8.
- Signor R, Ashbolt NJ (2009). Comparing probabilistic microbial risk assessments for drinking water against daily rather than annualised infection probability targets. J Water Health. 7(4):535–43.
- Signor RS, Roser DJ, Ashbolt NJ, Ball JE (2005). Quantifying the impact of runoff events on microbiological contaminant concentrations entering surface drinking source waters. J Water Health. 3(4):453–68.
- Sinton LW, Braithwaite RR, Hall CH, Mackenzie ML (2007). Survival of indicator and pathogenic bacteria in bovine feces on pasture. Appl Environ Microbiol. 73(24):7917–25.

- Smeets PW, van Dijk JC, Stanfield G, Rietveld LC, Medema GJ (2007). How can the UK statutory *Cryptosporidium* monitoring be used for quantitative risk assessment of *Cryptosporidium* in drinking water? J Water Health. 5(Suppl 1):107–18.
- Smeets PW, Rietveld LC, van Dijk JC, Medema GJ (2010). Practical applications of quantitative microbial risk assessment (QMRA) for water safety plans. Water Sci Technol. 61(6):1561–8.
- Sokolova E, Petterson SR, Dienus O, Nyström F, Lindgren P-E, Pettersson TJ (2015). Microbial risk assessment of drinking water based on hydrodynamic modelling of pathogen concentrations in source water. Sci Total Environ. 526:177–86.
- Soller JA, Eisenberg JN, Olivieri AW (1999). International Life Sciences Institute Research Foundation Risk Science Institute: evaluation of pathogen risk assessment framework. Oakland (CA): EOA Inc.
- Soller JA, Bartrand T, Ashbolt NJ, Ravenscroft J, Wade TJ (2010a). Estimating the primary etiologic agents in recreational freshwaters impacted by human sources of faecal contamination. Water Res. 44(16):4736–47.
- Soller JA, Schoen ME, Bartrand T, Ravenscroft JE, Ashbolt NJ (2010b). Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Res. 44(16):4674–91.
- Soller JA, Nellor MH, Cruz CJ, McDonald E (2015). Human health risk associated with direct potable reuse – evaluation through quantitative relative risk assessment case studies. Environ Sci Water Res Technol. 1:679–88.
- Steyn M, Jagals P, Genthe B (2004). Assessment of microbial infection risks posed by ingestion of water during domestic water use and full-contact recreation in a mid-southern African region. Water Sci Technol. 50(1):301–8.
- Stone DL, Harding AK, Hope BK, Slaughter-Mason S (2008). Exposure assessment and risk of gastrointestinal illness among surfers. J Toxicol Environ Health Assess. 71(24):1603–15.
- Suppes LM, Abrell L, Dufour AP, Reynolds KA (2014). Assessment of swimmer behaviors on pool water ingestion. J Water Health. 12(2):269–79.
- Symonds E, Verbyla M, Lukasik J, Kafle R, Breitbart M, Mihelcic J (2014). A case study of enteric virus removal and insights into the associated risk of water reuse for two wastewater treatment pond systems in Bolivia. Water Res. 65:257–70.
- Tadesse D, Desta A, Geyid A, Girma W, Fisseha S, Schmoll O (2010). Rapid assessment of drinking-water quality in the Federal Republic of Ethiopia. Country report for the pilot project implementation in 2004–2005. Geneva: World Health Organization and United Nations Children's Fund (UNICEF).
- Teunis PFM, Havelaar AH (1999). *Cryptosporidium* in drinking waters: evaluation of the ILSI/RSI quantitative risk assessment framework. Bilthoven: National Institute of Public Health and the Environment; 97 (RIVM Report 284 550 006).
- Teunis PFM, Havelaar A (2000). The Beta Poisson dose–response model is not a single hit model. Risk Anal. 20(4):511–8.
- Teunis PFM, Evers EG, Slob W (1999). Analysis of variable fractions resulting from microbial counts. Quant Microbiol. 1:63–88.
- Teunis PFM, Medema GJ, Kruidenier L, Havelaar AH (1997). Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. Water Res. 31(6):1333–46.
- Tseng LY, Jiang SC (2012). Comparison of recreational health risks associated with surfing and swimming in dry weather and post-storm conditions at southern California beaches using quantitative microbial risk assessment (QMRA). Mar Pollut Bull. 64(5):912–8.
- USDA/USEPA (2012). Microbial risk assessment guideline: pathogenic microorganisms with focus on food and water. Prepared by the Interagency Microbial Risk Assessment Guideline Workgroup. United States Department of Agriculture/ Food Safety and Inspection Service and United States Environmental Protection Agency (EPA/100/J-12/001; USDA/FSIS/ 2012-001; http://www.fsis.usda.gov/wps/wcm/connect/d79eaa29-c53a-451e-ba1c-36a76a6c6434/Microbial_Risk_ Assessment_Guideline_2012-001.pdf?MOD=AJPERES, accessed 9 April 2016).
- USEPA (2000). Estimated per capita water ingestion in the United States. Washington (DC): United States Environmental Protection Agency, Office of Water.
- USEPA (2003). Framework for cumulative risk assessment. Washington (DC): United States Environmental Protection Agency, National Center for Environmental Assessment (EPA/600/P-02/001F).
- USEPA (2004). Risk assessment principles and practices. Staff paper. Washington (DC): United States Environmental Protection Agency, Office of the Science Advisor (EPA/100/B-04/001).
- USEPA (2005). Economic analysis for the Final Long Term 2 Enhanced Surface Water Treatment Rule. Washington (DC): United States Environmental Protection Agency, Office of Water (http://nepis.epa.gov/Exe/ZyPDF.cgi/901S0000.PDF? Dockey=901S0000.PDF, accessed 19 April 2016).
- van den Akker B, Whiffin V, Cox P, Beatson P, Ashbolt NJ, Roser DJ (2011). Estimating the risk from sewage treatment plant effluent in the Sydney catchment area. Water Sci Technol. 63(8):1707–15.
- van der Sluijs J (2005). Uncertainty as a monster in the science policy interface: four coping strategies. Water Sci Technol. 52(6):87–92.

- van Heerden J, Ehlers MM, Vivier JC, Grabow WO (2005). Risk assessment of adenoviruses detected in treated drinking water and recreational water. J Appl Microbiol. 99(4):926–33.
- Wang X, Hunter PR (2010). A systematic review and meta-analysis of the association between self-reported diarrhoeal disease and distance from home to water source. Am J Trop Med Hyg. 83(3):582–4.
- Weir MH, Razzolini MRP, Rose JB, Masago Y (2011). Water reclamation redesign for reducing *Cryptosporidium* risks at a recreational spray park using stochastic models. Water Res. 45:6505–14.
- Westrell T (2004). Microbial risk assessment and its implications for risk management in urban water systems [PhD dissertation]. Linköping: University of Linköping, Faculty of Arts and Science.
- Westrell T, Bergstedt O, Stenström TA, Ashbolt NJ (2003). A theoretical approach to assess microbial risks due to failures in drinking water treatment. Int J Environ Health Res. 13(2):181–97.
- Westrell T, Schönning C, Stenström TA, Ashbolt NJ (2004). QMRA (quantitative microbial risk assessment) and HACCP (hazard analysis and critical control points) for management of pathogens in wastewater and sewage sludge treatment and reuse. Water Sci Technol. 50(2):23–30.
- WHO (1997). Guidelines for drinking-water quality, second edition. Vol. 3. Surveillance and control of community supplies. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/small-water-supplies-guidelines/en/, accessed 9 April 2016).
- WHO (2003). Guidelines for safe recreational water environments. Vol. 1. Coastal and fresh waters. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/srwe1/en/, accessed 9 April 2016).
- WHO (2006a). Guidelines for the safe use of wastewater, excreta and greywater. Vol. 2. Wastewater use in agriculture. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/gsuweg2/en/, accessed 9 April 2016).
- WHO (2006b). Guidelines for the safe use of wastewater, excreta and greywater. Vol. 4. Excreta and greywater use in agriculture. Geneva, World Health Organization (http://www.who.int/water_sanitation_health/publications/gsuweg4/en/, accessed 9 April 2016).
- WHO (2010). Human health risk assessment toolkit: chemical hazards. Geneva: World Health Organization (Harmonization Project Document No. 8; http://www.who.int/ipcs/methods/harmonization/areas/ra_toolkit/en/, accessed 9 April 2016).
- WHO (2011). Evaluating household water treatment options: health-based targets and microbiological performance specifications. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/household_water/en/, accessed 9 April 2016).
- WHO (2012). Water safety planning for small community water supplies. Step-by-step risk management guidance for drinking-water supplies in small communities. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/small-comm-water_supplies/en/, accessed 9 April 2016).
- WHO (2014a). Preventing diarrhoea through better water, sanitation and hygiene. Exposures and impacts in low- and middle-income countries. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/preventing-diarrhoea/en/, accessed 9 April 2016).
- WHO (2014b). Water safety plan: a field guide to improving drinking-water safety in small communities. Copenhagen: World Health Organization, Regional Office for Europe (http://www.euro.who.int/en/publications/abstracts/water-safety-plan-a-field-guide-to-improving-drinking-water-safety-in-small-communities, accessed 9 April 2016).
- WHO (2015a). Sanitation safety planning: manual for safe use and disposal of wastewater, greywater and excreta. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/ssp-manual/en/, accessed 9 April 2016).
- WHO (2015b). WHO estimates of the global burden of foodborne diseases. Foodborne diseases burden epidemiology reference group 2007–2015. Geneva: World Health Organization (http://www.who.int/foodsafety/publications/foodborne_disease/fergreport/en/, accessed 1 May 2016).
- WHO (in preparation). Guidelines for drinking-water quality, fourth edition incorporating first addendum. Geneva: World Health Organization.
- Williams BL, Florez Y, Pettygrove S (2001). Inter- and intra-ethnic variation in water intake, contact and source estimates among Tucson residents: implications for exposure analysis. J Expo Anal Environ Epidemiol. 11(6):510–21.
- Yang J, Schneider OD, Jjemba PK, LeChevallier MW (2015). Microbial risk modeling for main breaks. J Am Water Works Assoc. 107(2):E97–108.
- Yates MV, Gerba CP (1998). Microbial considerations in wastewater reclamation and reuse. In: Asano T, editor. Wastewater reclamation and reuse. Vol. 10. Lancaster (PA): Technomic Publishing Co. Inc.; 437–88.

ANNEX A | CASE-STUDIES

A1 Case-study 1: Pathogen risk to swimmers at non-sewage-impacted recreational beaches in the USA (Schoen & Ashbolt, 2010)

A1.1 Introduction

The USEPA uses faecal indicator bacterial criteria, based on epidemiological studies from human wastewater–impacted beaches, to identify bathing waters that exceed thresholds of concern (USEPA, 2012). There is concern, however, that non-sewage-impacted beaches may receive a different range and magnitude of pathogens from animal origin and may require an alternative standard to be equally protective of human health. Hence, the faecal indicator criteria may represent different levels of risk for beaches not impacted by human wastewaters. The high cost and impracticability of undertaking many epidemiological studies, along with the lack of any clear relationship between indicator and health outcome for non-human-impacted beaches (Boehm et al., 2009), led the USEPA to develop QMRA assessments for bathing beaches that are contaminated by different sources. The USEPA used QMRA to bring together scientific information regarding the relative occurrence of pathogens in different faecal sources and the infectivity and probability of illness associated with those pathogens to explore the relative contribution of different faecal sources to swimmers' health risk. The specific aim of this case-study was to explore how beaches contaminated by human sewage and by gulls may affect swimmers' health risk.

This case-study shows how QMRA can be used to derive site-specific bathing water criteria that are scientifically defensible and protective of human health.

A1.2 Problem formulation

The purpose of the QMRA was to:

- explore the risk to bathers associated with enterococci (i.e. faecal load) from seagulls compared with sewage for a water body at the USEPA-recommended water quality limit of culturable enterococci equal to 35 CFU·100 mL^{-1}; and
- predict when or if a non-sewage source of pathogens (seagulls) may dominate the illness risk in a water body affected by a mixture of sources.

The scope was defined by the following:

- **Hazard identification:** The hazards of concern were those causing gastrointestinal illness, sourced from poorly treated sewage and from gull excreta. Seagulls were assumed to contribute two reference pathogens: *Campylobacter jejuni* and *Salmonella enterica*. Sewage was assumed to contribute four reference pathogens commonly reported in wastewater effluent: norovirus, *Giardia intestinalis*, *Cryptosporidium* spp. and *S. enterica.*

Annexes

- **Exposure pathway:** Unintentional ingestion of surface waters while swimming in waters with enterococci at a concentration of 35 CFU·100 mL^{-1}, where the enterococci originated from either seagull excreta or poorly treated sewage (or a combination of the two).
- **Health outcome:** Probability of gastrointestinal illness.

A1.3 Exposure assessment

The exposure pathway and conceptual model applied to predicting exposure are illustrated in Fig. A1.1.

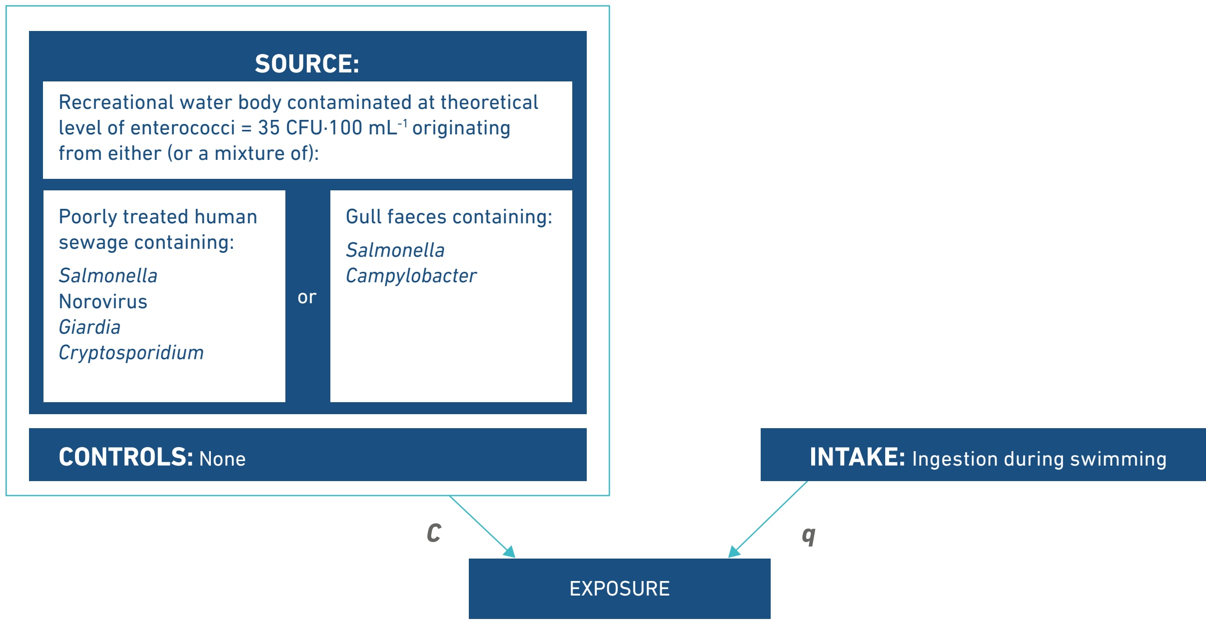


Fig. A1.1 Schematic of exposure pathway applied for comparing the exposure to pathogens depending on the faecal source to recreational waters (*C*: concentration of pathogens, microorganisms·L^{-1}; *q*: volume of water consumed, L) (adapted from Petterson & Ashbolt, 2015)

A1.3.1 Source

The exposure dose was estimated based on the enterococci concentration, with consideration given to the source of the faecal indicator relative to the pathogens, using the formula:

$$\mu_{rp}^{S} = \frac{C_{ENT} \cdot F^{S}}{R_{ENT}^{S} \cdot 100} \cdot R_{rp}^{S} \cdot p^{S} \cdot V \qquad \text{Eq. A1.1}$$

where:

μ_{rp}^{S} is the dose of each reference pathogen from each source (sewage or seagulls) (CFU, (oo)cysts or genomes)

C_{ENT} is the surf zone enterococci concentration (CFU·100 mL^{-1})

F^{S} is the fraction of enterococci from source *S*

R_{ENT}^{S} is the ratio of the count of enterococci to the wet mass of a composite seagull excreta sample (CFU·g^{-1}) or to the volume of sewage (CFU·L^{-1})

R_{rp}^{S} is the ratio of the count of reference pathogen to the wet mass of a composite gull excreta sample or to the volume of sewage

p^{S} is the fraction of human-infectious pathogenic strains for the reference pathogen from source *S*

V is the volume of water ingested (mL)

The quantitative distribution of values applied in the exposure assessment are summarized in Table A1.1. For full details of the quantification of model components and substantiation from the literature, the reader is referred to the original paper.

Table A1.1 Model input values for calculation of exposure dose

	Sewage		Gull	
	Distribution	**Reference**	**Distribution**	**Reference**
Concentration in faecal source material				
Enterococci ($CFU \cdot L^{-1}$ or g^{-1})	Uniform (10^{7}, 10^{8})	Tchobanoglous, Burton & Stensel (2003)	Uniform (10^{6}, 10^{8})	Fogarty et al. (2003); Haack, Fogarty & Wright (2003)
Campylobacter ($CFU \cdot L^{-1}$ or g^{-1})	NA		Uniform ($10^{3.3}$, $10^{6.0}$)	Levesque et al. (2000)
Salmonella ($CFU \cdot L^{-1}$ or g^{-1})	Uniform ($10^{0.5}$, $10^{3.0}$)	Lemarchand & Lebaron (2003)	Uniform ($10^{2.3}$, $10^{9.0}$)	Levesque et al. (2000)
Cryptosporidium ($oocysts \cdot L^{-1}$ or g^{-1})	Uniform ($10^{-0.3}$, $10^{4.6}$)	Rose et al. (2004)	NA	
Giardia ($cysts \cdot L^{-1}$ or g^{-1})	Uniform ($10^{0.8}$, $10^{4.0}$)	Rose et al. (2004)	NA	
Norovirus ($genomes \cdot L^{-1}$ or g^{-1})	Uniform ($10^{3.0}$, $10^{7.5}$)	Haramoto et al. (2006); Katayama et al. (2008)	NA	
Human-infectious fraction of pathogen strains				
All pathogens	1	NA	Uniform (0.01, 0.4)	Fenlon (1983)
Intake				
Exposure volume (mL)	Lognormal (2.92, 1.43)[a]			Dufour et al. (2006)

NA: not applicable

[a] Results for the combined population of adults and children.

Source: adapted with permission from Schoen & Ashbolt, 2010. Copyright 2010 American Chemical Society.

A1.3.2 Controls

None

A1.3.3 Intake

The volume of water ingested while swimming was based on the results of Dufour et al. (2006) for the combined population of adults and children.

A1.4 Health effects assessment

The dose–response relationships and the probability of illness given infection used in the study are summarized in Table A1.2.

A1.5 Risk characterization

Variability in the exposure parameters was modelled using Monte Carlo simulation of the defined distributions (Table A1.1). Point estimates of parameter values were used for dose–response models (Table A1.2). Therefore, the distribution of gastrointestinal illness risk captured the variability and uncertainty in pathogen dose (based on the variability in defined model inputs) and did not incorporate the uncertainty from the dose–response relationships. The influence of the selected dose–response models on the predicted risk was discussed in the original publication.

The predicted illness risk by pathogen is illustrated in Fig. A1.2, with the median illness risk associated with human sewage approximately 2 orders of magnitude higher than that associated with seagulls, illustrating that a water body at the recreational water quality limit may present a different risk to swimmers depending on the

source of the faecal contamination. The risk associated with the human sewage–impacted beaches was dominated by norovirus.

Table A1.2 Dose–response relationships and probability of illness given infection applied in the model

Reference pathogen	Dose units	Dose–response model	Probability of illness conditional on infection	Reference
Campylobacter jejuni				
Adults	CFU	Beta-Poisson $\alpha = 0.145$; $\beta = 7.59$	0.2	Medema et al. (1996)
Mixed population	CFU	Exact Beta-Poisson $\alpha = 0.024$; $\beta = 0.011$	$P_{ill\|inf} = 1 - (1 + \eta\mu)^{-r}$ $\eta = 10^{-8.44}$; $r = 10^{8.38}$	Teunis et al. (2005)
Salmonella enterica serotype Bareilly	CFU	Gompertz model $\ln(a) = 11.68$, $b = 0.82$	NA: model predicts illness directly from dose	Coleman & Marks (1998, 2000); Coleman et al. (2004)
Cryptosporidium parvum	Oocysts	Exponential $r = 0.09$	0.7	USEPA (2005)
Giardia intestinalis	Cysts	Exponential $r = 0.019\ 9$	0.9	Rose, Haas & Regli (1991)
Norovirus	Genomes	Exact Beta-Poisson including parameter for clumping $\alpha = 0.04$; $\beta = 0.055$; $a = 0.000\ 1$	0.68	Teunis et al. (2008)

NA: not applicable

Source: adapted with permission from Schoen & Ashbolt, 2010. Copyright 2010 American Chemical Society.

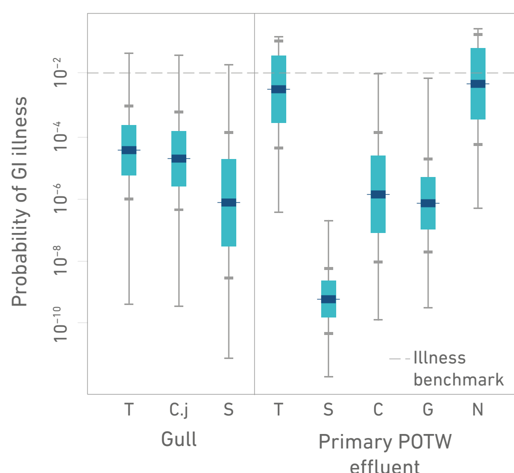


Fig. A1.2 Predicted gastrointestinal (GI) illness by reference pathogen (median, interquartile range, 10th and 90th percentiles, minimum and maximum) for adults following a single accidental ingestion of recreational water containing fresh faecal contamination at 35 CFU·100 mL^{-1} enterococci contributed by seagulls (Gull) or primary publicly owned wastewater treatment works (POTW) effluent: T, total gastrointestinal risk; C.j, *C. jejuni* risk; S, *Salmonella* risk; C, *Cryptosporidium* risk; G, *Giardia* risk; N, norovirus risk (Adapted with permission from Schoen & Ashbolt, 2010. Copyright 2010 American Chemical Society)

Sensitivity analysis provided a useful tool for exploring the interaction between model parameters. The authors first undertook a sensitivity analysis of the parameters driving exposure using the Spearman rank correlation coefficients of the Monte Carlo simulation (Fig. A1.3). For the risk attributable to seagulls, the volume of water ingested (V), the density of faecal indicator bacteria (enterococci) in faeces and the density of *C. jejuni* in faeces were of relative equivalent importance. The density of *Salmonella* in the faeces and the fraction of human-infectious pathogens in the faeces were relatively less important.

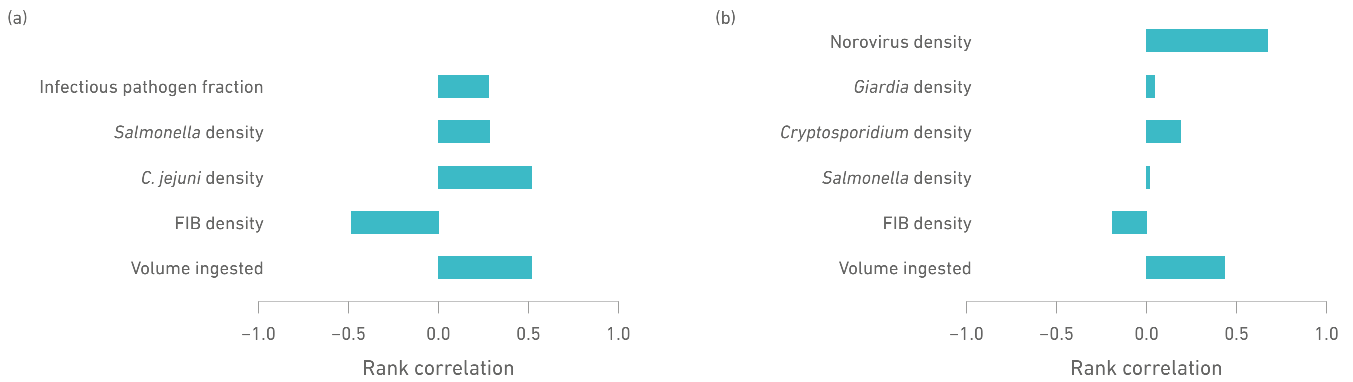


Fig. A1.3 Spearman rank correlation coefficient for exposure parameter inputs to the predicted probability of gastrointestinal illness from accidental ingestion of recreational water containing fresh faecal contamination at 35 CFU·100 mL^{-1} enterococci for (a) seagull excreta and (b) primary wastewater effluent. FIB: faecal indicator bacteria (Adapted with permission from Schoen & Ashbolt, 2010. Copyright 2010 American Chemical Society).

Owing to the lack of data, the authors were specifically concerned with further investigating the sensitivity of the calculated risk to the fraction of human-infectious pathogen strains from gulls. The results of the parametric sensitivity analysis are illustrated in Fig. A1.4. The illness risk associated with gull excreta further decreased as the assumed fraction of the total pathogen strains from gulls that are human infectious decreased below the best estimate of 0.2.

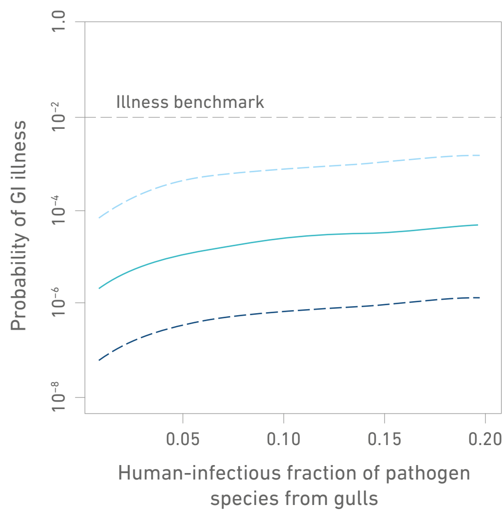


Fig. A1.4 Parametric sensitivity analysis of the predicted probability of gastrointestinal (GI) illness (median, 10th and 90th percentiles) for adults attributable to *Campylobacter jejuni* from accidental ingestion of recreational water containing fresh seagull excreta contamination at 35 CFU·100 mL^{-1} enterococci to changes in the assumed fraction of total *C. jejuni* strains from seagulls that are infectious to humans (Adapted with permission from Schoen & Ashbolt, 2010. Copyright 2010 American Chemical Society)

The model was also used to explore the relative importance of the two sources when there is a mixture (Fig. A1.5). Considering all possible mixtures along the x-axis, the median predicted probability of gastrointestinal illness from gulls was greater than that from a sewage mixture only when gulls represented greater than 98% of the faecal indicator load. Therefore, the dominant source of faecal indicators at a recreational beach may not be the dominant source of risk. Conversely, a little fresh sewage contamination mixed with non-sewage source(s) may dominate risk, and possibly at a level that may exceed the benchmark risk, when more than 20% of enterococci originate from sewage.

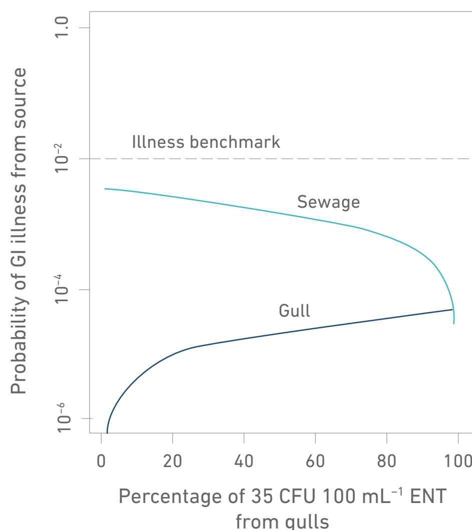


Fig. A1.5 Comparison of median gastrointestinal (GI) illness risk for adults when enterococci concentration (ENT at 35 CFU·100 mL^{-1}) is attributed to a mixture of primary sewage effluent and seagull excreta (Adapted with permission from Schoen & Ashbolt, 2010. Copyright 2010 American Chemical Society)

A1.6 Risk management

The results of the QMRA provide a quantitative basis to inform the understanding of the dynamics associated with pathogen risks from different sources of contamination. It is theoretically considered that the source of the faecal indicators will influence the consumer risk, and in this study, the dynamics of different sources were quantitatively explored. Firstly, the results confirm, as expected, that, at the same concentration of enterococci, the predicted risks to swimmers were lower in waters contaminated by gulls than in waters contaminated by sewage. Secondly, the study identified that the dominant source of faecal indicators may not be the dominant source of risk. Site evaluation by sanitary survey and MST approaches cannot simply focus on prioritizing faecal sources based on size, but must recognize that smaller sources may be important to the downstream risk profile. The USEPA has used this and similar studies to develop a QMRA-based approach to derive alternative recreational water criteria that are equally protective of human health in non-sewage-impacted beaches.

A1.7 Evaluation of the QMRA

This study allowed for the quantitative issues associated with faecal indicator sources to be explored and provided a quantitative basis for the assumption that non-human-impacted waters may pose a lower risk to consumers than human-impacted waters, an issue that could not be easily explored using epidemiological approaches alone. Whereas many of the model inputs were uncertain, sensitivity analysis was applied to understand the dynamics and inter-relationships between model variables, which proved valuable. The QMRA was considered a strong tool to compare the risk of alternative pathogen sources and was used by the USEPA to evaluate agricultural sources as well as to inform the near ambient water quality criteria (USEPA, 2012). The study also facilitated the identification of research and data gaps, such as the importance of quantifying the probability that a pathogen from different animal sources is human infectious.

A1.8 References

- Boehm AB, Ashbolt NJ, Colford JM Jr, Dunbar LE, Fleming LE, Gold MA et al. (2009). A sea change ahead for recreational water quality criteria. J Water Health. 7(1):9–20.
- Coleman M, Marks H (1998). Topics in dose–response modeling. J Food Prot. 61(11):1550–9.
- Coleman ME, Marks HM (2000). Mechanistic modeling of salmonellosis. Quant Microbiol. 2:227–47.
- Coleman ME, Marks HM, Golden NJ, Latimer HK (2004). Discerning strain effects in microbial dose–response data. J Toxicol Environ Health Assess. 67(8–10):667–85.
- Dufour AP, Evans O, Behymer TD, Cantú R (2006). Water ingestion during swimming activities in a pool: a pilot study. J Water Health. 4(4):425–30.
- Fenlon DR (1983). A comparison of *Salmonella* serotypes found in the faeces of gulls feeding at a sewage works with serotypes present in the sewage. J Hyg (Lond). 91(1):47–52.

- Fogarty LR, Haack SK, Wolcott MJ, Whitman RL (2003). Abundance and characteristics of the recreational water quality indicator bacteria *Escherichia coli* and enterococci in gull faeces. J Appl Microbiol. 94(5):865–78.
- Haack SK, Fogarty LR, Wright C (2003). *Escherichia coli* and enterococci at beaches in the Grand Traverse Bay, Lake Michigan: sources, characteristics, and environmental pathways. Environ Sci Technol. 37(15):3275–82.
- Haramoto E, Katayama H, Oguma K, Yamashita H, Tajima A, Nakajima H et al. (2006). Seasonal profiles of human noroviruses and indicator bacteria in a wastewater treatment plant in Tokyo, Japan. Water Sci Technol. 54(11–12):301–8.
- Katayama H, Haramoto E, Oguma K, Yamashita H, Tajima A, Nakajima H et al. (2008). One-year monthly quantitative survey of noroviruses, enteroviruses, and adenoviruses in wastewater collected from six plants in Japan. Water Res. 42(6–7):1441–8.
- Lemarchand K, Lebaron P (2003). Occurrence of *Salmonella* spp and *Cryptosporidium* spp in a French coastal watershed: relationship with fecal indicators. FEMS Microbiol Lett. 218(1):203–9.
- Levesque B, Brousseau P, Bernier F, Dewailly E, Joly J (2000). Study of the bacterial content of ring-billed gull droppings in relation to recreational water quality. Water Res. 34(4):1089–96.
- Medema GJ, Teunis PF, Havelaar AH, Haas CN (1996). Assessment of the dose–response relationship of *Campylobacter jejuni*. Int J Food Microbiol. 30(1–2):101–11.
- Petterson SR, Ashbolt NJ (2015). Exposure assessment. In: Yates MV, Nakatsu CH, Miller RV, Pillai SD, editors. Manual of environmental microbiology, fourth edition. ASM Press.
- Rose JB, Haas CN, Regli S (1991). Risk assessment and control of waterborne giardiasis. Am J Public Health. 81:709–13.
- Rose JB, Nowlin H, Farrah SR, Harwood V, Levine AD, Lukasik J et al. (2004). Reduction of pathogens, indicator bacteria, and alternative indicators by wastewater treatment and reclamation processes. Water Environment Research Foundation (WERF Report: Water for Reuse (00-PUM-2T)).
- Schoen ME, Ashbolt NJ (2010). Assessing pathogen risk to swimmers at non-sewage impacted recreational beaches. Environ Sci Technol. 44(7):2286–91.
- Tchobanoglous G, Burton FL, Stensel HD (2003). Wastewater engineering: treatment and reuse, fourth edition. New York (NY): Metcalf & Eddy, Inc.
- Teunis P, van den Brandhof W, Nauta M, Wagenaar J, van den Kerkhof H, van Pelt W (2005). A reconsideration of *Campylobacter* dose–response relation. Epidemiol Infect. 133:583–92.
- Teunis PFM, Moe CL, Liu P, Miller SE, Lindesmith L, Barie RS et al. (2008). Norwalk virus: how infectious is it? J Med Virol. 80:1468–76.
- USEPA (2005). Economic analysis for the Final Long Term 2 Enhanced Surface Water Treatment Rule. Washington (DC): United States Environmental Protection Agency, Office of Water (http://nepis.epa.gov/Exe/ZyPDF.cgi/901S0000.PDF?Dockey=901S0000.PDF, accessed 19 April 2016).
- USEPA (2012). Recreational water quality criteria. Washington (DC): United States Environmental Protection Agency, Office of Water (EPA Report 820-F-12-058).

A2 Case-study 2: Water reclamation redesign for reducing *Cryptosporidium* risks at a recreational spray park in the USA (Weir et al., 2011)

A2.1 Introduction

Outbreaks associated with spray parks, also known as recreational waterparks, are well recognized and may be partly due to the engineering designs used for their water reclamation systems, which are problematic to control. Weir and coworkers (2011) undertook this investigation based on an outbreak of cryptosporidiosis linked to a spray park in New York State, where it had been determined by epidemiological investigation that the spray pad (the main attraction) was the primary exposure point. The outbreak affected 746 people who reported gastrointestinal illness after visiting the spray park facility.

After the outbreak, an investigation and risk assessment were launched to determine the likely cause and to evaluate potential design changes to prevent future outbreaks. This case-study illustrates the use of QMRA to identify different management strategies that could reduce future outbreaks from spray parks.

A2.2 Problem formulation

The purpose of the QMRA was to evaluate possible gastrointestinal risk due to exposure to water contaminated with *Cryptosporidium* oocysts and to evaluate the efficacy of potential combined treatment retrofits aimed at reducing recreational risks. Two main design changes were proposed: removing the pipe between Tank 2 and Tank 1 (which was likely acting as a bypass) from the treatment system, and the combination of removing this pipe and adding an ozone contactor.

The scope was defined by the following:

- **Hazard identification:** The assessment focused on *Cryptosporidium* oocysts, the source of the outbreak in 2005.
- **Exposure pathways:** The source of *Cryptosporidium* oocysts was considered to be associated with oocyst-infected individuals contaminating the water during recreation. The contaminated water was then assumed to be recirculated through the system, and individuals were exposed by involuntary consumption during recreational activities.
- **Health outcomes:** Probability of infection and probability of illness.

A2.3 Exposure assessment

The exposure was assessed for a 3-hour simulation of the spray pad system following a hypothetical faecal release incident. The exposure pathway modelled in the QMRA is illustrated in Fig. A2.1.

A2.3.1 Source

A hypothetical faecal release of 2×10^7 oocysts (approximately 2 g of faeces) (Chappell et al., 1996; Yoder & Beach, 2007) was simulated.

A2.3.2 Controls

- **Dilution:** The treatment system received recreational water from two drains in the centre of the spray pad (Fig. A2.2a is a flow chart of the unaltered system). Water is retained from the spray pad in Tank 1, before being treated by filtration and free chlorine disinfection. Tank 1 is also used to hold and return treated water to the spray pad, which has the unintended effect of bypassing the treatment system.
- **Inactivation:** The inactivation of oocysts in neutral water was assumed to be $0.000\ 363 \cdot h^{-1}$ (Robertson, Campbell & Smith, 1992).
- **Filtration:** The filter was assumed to operate continuously and properly during the season with a sufficient schmutzdecke (complex biological scum layer formed at the top of a filter, essentially the main treatment layer of the filter) developed for removal of oocysts (and other pathogens). Treatment efficacy = 0.99 (Schuler, Ghosh & Gopalan, 1991).
- **Ozone:** The oocyst inactivation fraction at $2\ mg \cdot L^{-1}$ for 1 minute was assumed to be 0.79 (Corona-Vasquez et al., 2002).

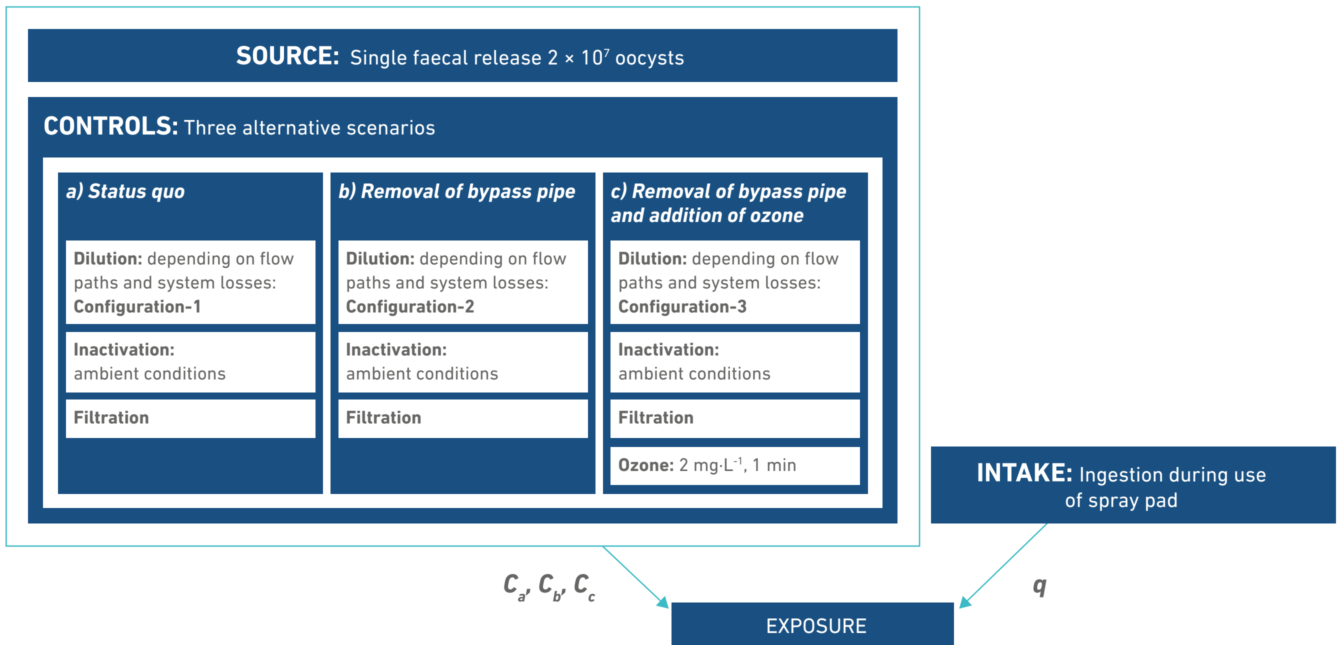


Fig. A2.1 Schematic of exposure pathway applied for assessing user risks under current conditions: a) status quo; and for two proposed retrofits: b) removal of the bypass pipe and c) removal of the bypass pipe combined with the addition of an ozone contactor

For complete details of flow assumptions and model parameterization, the reader should refer to the original publication and supplementary material.

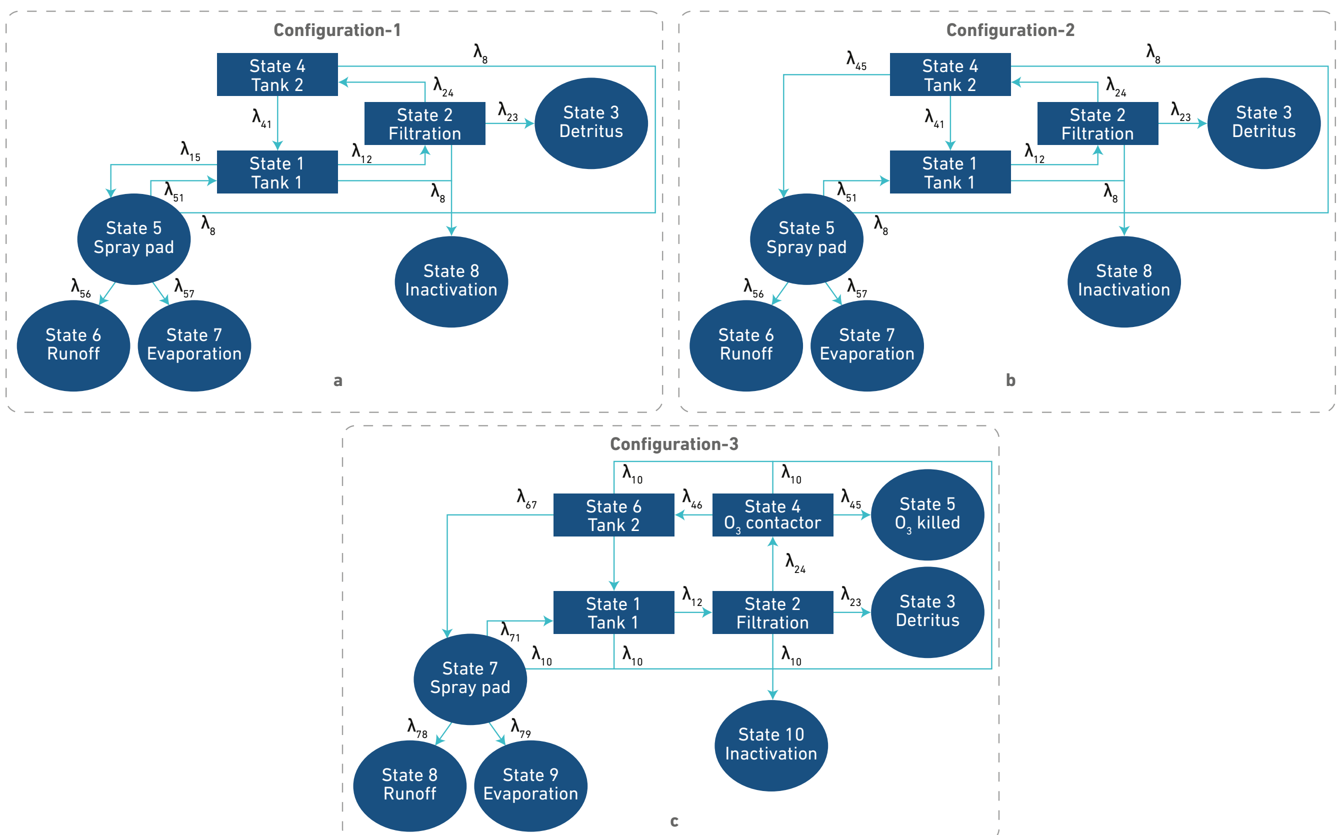


Fig. A2.2 Markov chain model flow charts of system configurations: a) configuration-1, which is the current unchanged system; b) configuration-2, which is the first suggested change, with the removal of the bypass pipe; and c) configuration-3, which is the last recommended change, where the bypass pipe is removed and an ozone contactor is included as well (adapted from Weir et al., 2011)

Given the complexity of water flow within the reclamation and treatment system, a Markov chain model was constructed to simulate the flows and efficacy through the system. A distribution was fitted to the simulated concentration of oocysts in the water released at the spray pad under the existing configuration (C_a) (configuration-1) and with the removal of the bypass pipe (C_b) (configuration-2) and the removal of the bypass pipe with inclusion of ozone (C_c) (configuration-3).

A2.3.3 Intake

The exposure volume was defined using data on accidental ingestion of water during swimming from Dufour et al. (2006) as an assumed constant value of 0.108 L.

A2.4 Health effects assessment

The dose–response relationship for probability of infection due to ingestion of *Cryptosporidium* oocysts was described by the exponential dose–response model (i.e. a single-parameter model, referred to in the original publication as *k*, here identified as *r* for consistency). The parameter used in the analysis was optimized from human volunteer studies (USEPA, 2005b); *r* was defined by a triangular distribution (0.0074, 0.0907, 0.3044) reflecting minimal, average and maximum likely values for *r*. The probability of illness given infection was assumed to be 0.5, citing USEPA (2005a).

A2.5 Risk characterization

Monte Carlo simulation was used to characterize the risk (P_{inf} and P_{ill}) to the population exposed to water contaminated with *Cryptosporidium* oocysts with the current treatment system and under each of the two modifications. The results are illustrated in Fig. A2.3. For the current system, 60% of the infection risks were equal to or greater than 0.9. When the bypass was eliminated from the water reclamation system and all the water was filtered, the overall risk levels were reduced (with 44% of the infection risks equal to or greater than 0.9). When the bypass was removed and an ozone contactor was included, the risk levels were further reduced (with 32% of the infection risks equal to or greater than 0.9).

A2.6 Risk management

In response to the results of the risk analysis, it was recommended that, at a minimum, the removal of the bypass pipe would be necessary to reduce the risks associated with waterborne disease. Although the inclusion of the ozone contactor would be expected to further reduce the risks, this was not at a sufficiently appreciable risk reduction compared with the less costly retrofit.

The authors also identified that, given the high relative risks still related to this spray park and the exposure scenario, a risk communication strategy may be a simple and very cost-effective means of mitigating risk to users – something as simple as a pamphlet warning users of potential pathogen exposure or a staff member who can be consulted with questions that users may have regarding the potential risks from spray park recreation. This study illustrates that water treatment for recreational venues such as spray parks demands more attention to protect public health.

A2.7 Evaluation of the QMRA

The systematic approach of investigating an outbreak and undertaking a QMRA that included a stochastic model to simulate the treatment system brought to light the relatively high potential risk from the water reclamation system used at the spray park. Identification of two potential risk mitigation strategies from the problem formulation stage drove the scope of the risk assessment and allowed for the predicted risks associated with each scenario to be compared, providing quantitative input for decision support.

A2.8 References

- Chappell CL, Okhuysen PC, Sterling CR, DuPont HL (1996). *Cryptosporidium parvum*: intensity of infection and oocyst excretion patterns in healthy volunteers. J Infect Dis. 173(1):232–6.
- Corona-Vasquez B, Samuelson A, Rennecker JL, Mariñas BJ (2002). Inactivation of *Cryptosporidium parvum* oocysts with ozone and free chlorine. Water Res. 36(16):4053–63.

- Dufour AP, Evans O, Behymer TD, Cantú R (2006). Water ingestion during swimming activities in a pool: a pilot study. J Water Health. 4(4):425–30.
- Robertson LJ, Campbell AT, Smith HV (1992). Survival of *Cryptosporidium parvum* oocysts under various environmental pressures. Appl Environ Microbiol. 58(11):3494–500.
- Schuler PF, Ghosh MM, Gopalan P (1991). Slow sand diatomaceous earth filtration of cysts and other particulates. Water Res. 25:995–1006.
- USEPA (2005a). Economic analysis for the Final Long Term 2 Enhanced Surface Water Treatment Rule. Washington (DC): United States Environmental Protection Agency, Office of Water (EPA 815-R-06-001).
- USEPA (2005b). Appendices to: Economic analysis for the Final Long Term 2 Enhanced Surface Water Treatment Rule. Vol. II. Washington (DC): United States Environmental Protection Agency, Office of Water (EPA 815-R-06-001).
- Weir MH, Razzolini MRP, Rose JB, Masago Y (2011). Water reclamation redesign for reducing *Cryptosporidium* risks at a recreational spray park using stochastic models. Water Res. 45:6505–14.
- Yoder JS, Beach MJ (2007). Cryptosporidiosis surveillance – United States, 2003–2005. MMWR Surveill Summ. 56 (SS-07):1–10 (http://www.cdc.gov/mmwr/PDF/ss/ss5607.pdf, accessed 9 April 2016).

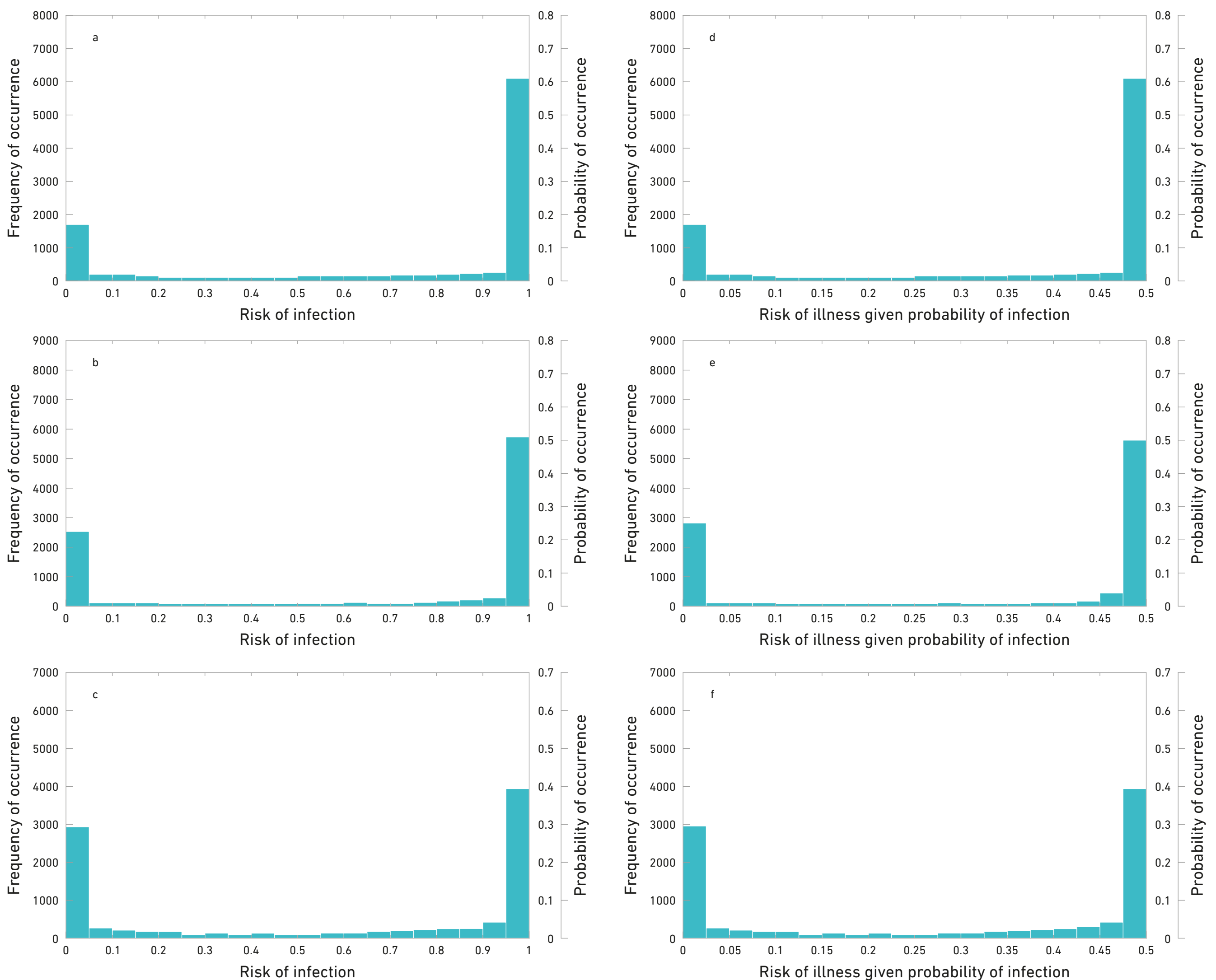


Fig. A2.3 Monte Carlo risk model results for risk of infection for a) configuration-1, b) configuration-2 and c) configuration-3; and for risk of illness for d) configuration-1, e) configuration-2 and f) configuration-3, where the x-axis is the risk level experienced during the simulation, the primary y-axis is the frequency of observing that risk level from the 10 000 iterations performance and the secondary y-axis is the probability that each risk level was encountered during the simulation (adapted from Weir et al., 2011)

A3 Case-study 3: Evaluating *Cryptosporidium* risk at a large number of drinking-water systems in France (Medema et al., 2009)

A3.1 Introduction

Suez Environnement operates, through Lyonnaise-des-Eaux, a large number (>1500) of drinking-water systems in France. After the waterborne outbreaks of cryptosporidiosis reported in the USA and the United Kingdom, Suez Environnement wanted to develop an approach to evaluate the risk of transmitting *Cryptosporidium* through their water systems. This case-study illustrates how QMRA allowed an objective, quantitative and transparent process to provide management with sufficient information to discriminate *Cryptosporidium* risk between many water systems and determine priorities for risk management.

A3.2 Problem formulation

The purpose of the QMRA was to:

- evaluate compliance with the European Union drinking-water directive (which states that drinking-water must be free from "parasites … which in numbers or concentrations, … constitute a potential danger to human health"; EU, 2014);
- identify whether any of these systems was at risk from *Cryptosporidium*; and
- prioritize investments (if needed).

The scope of the QMRA was defined by:

- **Hazard identification:** The assessment was undertaken only for the reference pathogen *Cryptosporidium*, because of its likely presence in source waters and the difficulty to remove or inactivate the pathogen by treatment, particularly as a result of its resistance to chemical disinfection.
- **Exposure pathways:** Surface water from a range of different catchment types and groundwater sources, treated by a range of different treatment trains, and exposure via the consumption of unboiled tap water.
- **Health outcomes:** Annual probability of infection.

A3.3 Exposure assessment

A questionnaire was sent to the operators of each system, inquiring about the volume of water produced, the type of source water used, including information on the type of environment (urban, rural, presence of cattle, etc.) and data on general water quality parameters (coliforms, turbidity, ammonium and nitrate), and the type of treatment processes used. The returned information covered treatment facilities that supply 9 million people with drinking-water.

A3.3.1 Source

The information from the survey was used to categorize the different water systems. These included groundwater systems, groundwater systems under the influence of surface water, surface water systems, and systems with drinking-water that were a blend of the former systems. Information on the occurrence of *Cryptosporidium* in surface water and groundwater was available from the scientific literature (Fig. A3.1). It was assumed that these data were representative for all systems, and they were used to estimate the concentration of *Cryptosporidium* in each of the source water types identified, using a conservative estimation (Fig. A3.2).

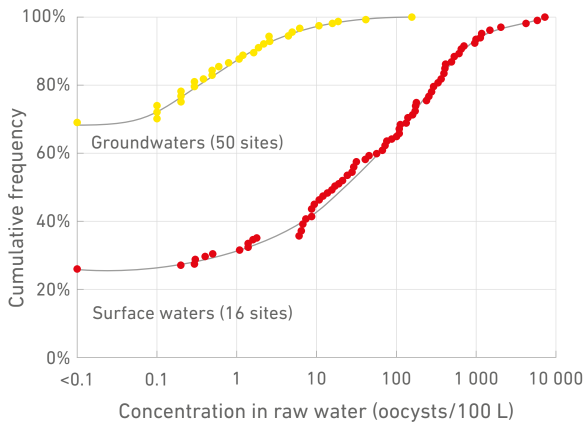


Fig. A3.1 Cumulative distribution of the concentration of *Cryptosporidium* in surface waters and in groundwaters (Medema et al., 2009)

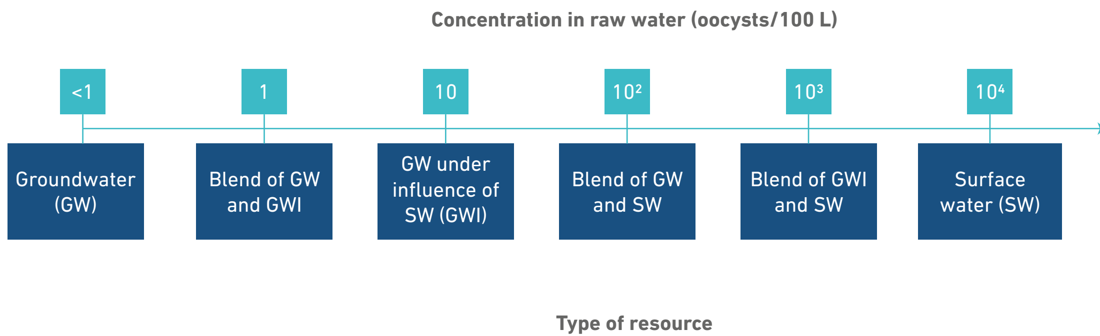


Fig. A3.2 Estimated concentration of *Cryptosporidium* in source waters, based on the type of source water (Medema et al., 2009)

A3.3.2 Controls

Information on the removal or inactivation of *Cryptosporidium* by various treatment processes was obtained from studies performed by the Centre International de Recherche sur l'Eau et l'Environnement (CIRSEE), the main research and expertise centre of Suez Environnement, and collected from the scientific literature. Generic $\log_{10}$ removals were assigned to each of the individual treatment processes (Fig. A3.3) under the assumptions that 1) the generic $\log_{10}$ removals described the treatment performance at each site adequately and 2) performance of treatment processes is constant and independent.

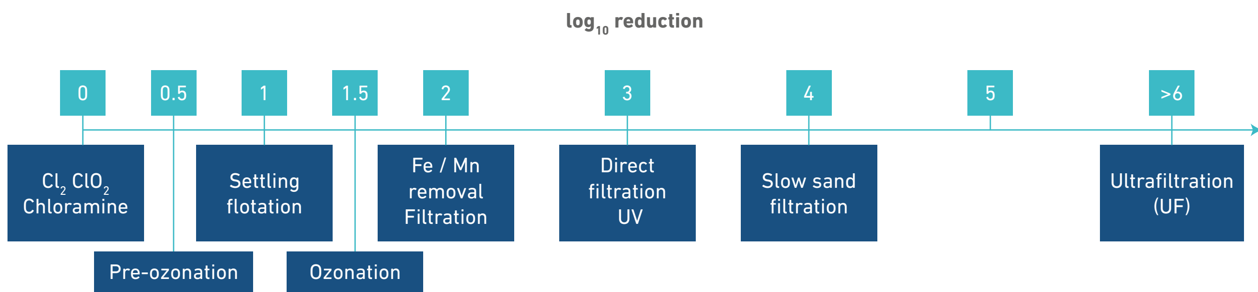


Fig. A3.3 Log_{10} removals for *Cryptosporidium* reduction by treatment processes (adapted from Medema et al., 2009)

The information about each treatment works was entered into a database with the facility to analyse the data.

A3.3.3 Intake

The average daily consumption of unheated tap water was assumed to be 1 L per person.

A3.4 Health effects assessment

The exponential dose–response model of *Cryptosporidium parvum* (Iowa strain; see Annex D) was used ($r = 0.004$) (DuPont et al., 1995).

A3.5 Risk characterization

For each system, a treatment performance target was defined according to the estimated concentration of *Cryptosporidium* in source water, in order to achieve one of these three levels of risk:

1) A low level of risk was related to a *Cryptosporidium* concentration of 0.003 per 100 L of source water. This concentration was derived from a health-based target of a probability of infection of 10^{-4} pppy (Haas et al., 1996).
2) A medium level of risk was related to the analytical detection limit of *Cryptosporidium*, which was determined to be 0.3 oocyst per 100 L. The associated probability of infection was 10^{-2} pppy.
3) A high level of risk was arbitrarily set at 30 oocysts per 100 L. The associated probability of infection was 0.6 pppy.

The treatment capacity of each treatment facility was compared with these treatment performance targets, and a level of risk was assigned to each facility accordingly (Fig. A3.4).

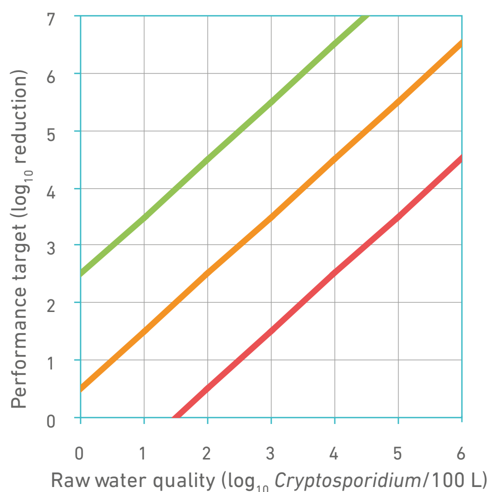


Fig. A3.4 Matrix to determine the level of risk associated with a treatment facility, based on the estimated concentration of *Cryptosporidium* in source water and the performance (log_{10} credits) of the treatment processes (adapted from Medema et al., 2009)

A3.6 Risk management

The risk assessment highlighted that the sites at higher risk were primarily small systems (serving <5000 people) and groundwater systems that were under the influence of surface water. The operating conditions of the treatment plants located in the medium- and high-risk categories were reviewed in order to check that the recommended best practices were being applied.

To support the outcome of the risk assessment, a *Cryptosporidium* monitoring programme was conducted at selected sites from each of the risk categories. In this monitoring programme, treated water samples were collected every 2 weeks for a period of at least 6 months and analysed for *Cryptosporidium* (and *Giardia*) and other water quality parameters: turbidity, coliforms, aerobic spores, temperature, pH, conductivity, UV transmission and ammonium. Additional samples were taken during turbidity peaks.

The results of the *Cryptosporidium* monitoring were consistent with the risk assessment: high-risk sites showed the highest frequency of samples positive for *Cryptosporidium*, and the highest concentrations were observed at these sites. No *Cryptosporidium* was found at any of the low-risk sites, and the medium-risk sites gave intermediate results. There was also a good correlation between the presence of *Cryptosporidium* and high turbidity. The verification of the QMRA using the monitoring data was very valuable to convince the risk managers and owners of the water supplies that the QMRA results are a valid and good basis for setting risk management priorities.

The main risk factors for *Cryptosporidium* that were found were the presence of cattle in the catchment area, less than 99% compliance with the coliform standard and turbidity of greater than 0.2 nephelometric turbidity unit (NTU) in distributed water. The company has audited the high-risk sites and, where necessary, has upgraded the treatment facilities in accordance with the local health authority. Furthermore, Suez Environnement has since used the same risk assessment approach for the waterworks that they operate in other countries.

A3.7 Evaluation of the QMRA

The QMRA provided adequate information for Suez Environnement to underpin the need for further risk control measures in some of their systems and to prioritize the systems. The QMRA was also a means to demonstrate due diligence to the customers and regulator. The QMRA guided the next steps: audit of the medium- and high-risk sites and confirmation of the conclusions by a dedicated *Cryptosporidium* monitoring programme.

System population size could have been added to inform the prioritization, as larger systems in the high-risk category have a greater potential overall public health impact compared with smaller systems. It is noted, however, that smaller systems may not be operated as well as larger systems. Furthermore, a sensitivity analysis, wherein each of the input values is replaced by a value that would represent an extreme event (e.g. peak contamination in source water, [partial] failure of the treatment process), could provide information about the robustness of the systems and the relative contribution of each of the steps in the system to overall safety.

A3.8 References

- DuPont HL, Chappell CL, Sterling CR, Okhuysen PC, Rose JB, Jakubowski W (1995). The infectivity of *Cryptosporidium parvum* in healthy volunteers. N Engl J Med. 332(13):855–9.
- EU (2014). European Union (Drinking Water) Regulations 2014. Statutory Instruments (S.I. No. 122 of 2014; https://www.fsai.ie/uploadedFiles/Legislation/Food_Legisation_Links/Water/SI122_2014.pdf, accessed 20 April 2016).
- Haas CN, Crockett CS, Rose JB, Gerba CP, Fazil AM (1996). Assessing the risk posed by oocysts in drinking water. J Am Water Works Assoc. 88(9):131–6.
- Medema GJ, Teunis PF, Blokker M, Deere D, Davidson D, Charles P et al. (2009). Risk assessment of *Cryptosporidium* in drinking water. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/cryptoRA/en/, accessed 9 April 2016).

A4 Case-study 4: USEPA Long Term Surface Water Treatment Rule – health benefit of a new drinking-water regulation in the USA (USEPA, 2005)

A4.1 Introduction

The Long Term 2 Enhanced Surface Water Treatment Rule (LT2ESWTR) was developed by the USEPA in 2005 to strengthen protection against microbial contaminants, especially *Cryptosporidium*, in drinking-water delivered by public water supplies that use surface water or groundwater under the direct influence of surface water. As a result, systems now monitor their public water for *Cryptosporidium* sources to provide further data to reassess against the standard treatment requirements set out in the LT2ESWTR for the control of *Cryptosporidium* via drinking-water.

This case-study addresses the USEPA's requirement to demonstrate the health costs and benefits of a new regulation, here described for *Cryptosporidium* management to control endemic cryptosporidiosis via drinking-water. It also provides an example of a probabilistic risk analysis that evaluates both variability and uncertainty at the same time and includes probabilistic treatments of uncertain dose–response and occurrence parameters.

A4.2 Problem formulation

The purpose of the QMRA was to estimate the number of illnesses and deaths associated with endemic cryptosporidiosis that would be avoided because of the LT2ESWTR. The quantified health benefits estimated were derived from calculating the difference between illness and death estimates for the baseline (pre-LT2ESWTR) condition and illness and death estimates after implementation of the LT2ESWTR. This rule was expected to reduce health effects associated with other pathogens such as *Giardia* and other waterborne bacterial or viral pathogens, although these health benefits were not quantified.

The scope was defined by:

- **Hazard identification:** *Cryptosporidium* was selected as a significant concern in drinking-water because it commonly occurs in surface waters, is resistant to chlorine-based disinfectants and causes numerous waterborne disease outbreaks. Consumption of water containing *Cryptosporidium* can cause gastrointestinal illness, which may be severe and sometimes fatal for people with weakened immune systems (which may include very young children, the elderly, and people with human immunodeficiency virus/acquired immunodeficiency syndrome [HIV/AIDS]).
- **Exposure pathways:** Drinking-water produced from surface water and groundwater under the direct influence of surface water systems.
- **Health outcomes:** Cases of cryptosporidiosis and associated deaths.

A4.3 Exposure assessment

A4.3.1 Source

To characterize the exposure to infectious *Cryptosporidium* oocysts in surface waters, the USEPA estimated the source water contamination levels using each of three data sets:

1) Information Collection Rule (ICR) for all large systems;
2) Information Collection Rule Supplemental Surveys Medium (ICRSSM) for systems serving 10 000–99 999 people; and
3) Information Collection Rule Supplemental Surveys Large (ICRSSL) for large plants serving more than 100 000 people.

A lognormal distribution was fitted to each data set using the number of oocysts detected (counts), the associated sample volume analysed and test method recovery rates and assuming that each observed count came from a Poisson probability distribution.

The percentage of oocysts that were infectious was modelled by a triangular distribution having a lower bound of 30%, a mode of 40% and an upper bound of 50% for data sets derived from the ICRSSM and ICRSSL (data sets 2 and 3) and a lower bound of 15%, a mode of 20% and an upper bound of 25% for data sets derived from the ICR (data set 1).

A4.3.2 Controls

The USEPA used the 2000 Community Water System Survey data to obtain information on the number of treatment plants per system, the source of water treated at each plant and the type of treatment in place. The analysis excluded systems whose flow rates per person were unusually high or low and those plants that treated groundwater or purchased water only. The predicted $\log_{10}$ removal achieved for systems was modelled using a triangular distribution (defined by minimum, mode and maximum) as follows:

- **For small systems (serving fewer than 10 000 people):** minimum to maximum range of 2–4 $\log_{10}$ removal to capture system-to-system variability. Mode varied between 2.25 and 3.25 $\log_{10}$ removal, intended to capture uncertainty.
- **For large systems (those serving at least 10 000 people):** minimum to maximum range of 2–5 $\log_{10}$ removal to capture system-to-system variability. Mode varied between 2.5 and 3.5 $\log_{10}$ removal, intended to capture uncertainty.

A4.3.3 Intake

The volume of individual daily drinking-water consumption was estimated to be 1.07 L per person per day. To account for the annual days of exposure, the model assigned exposure days per year for community water systems at 350, for non-transient non-community water systems at 250 and for transient non-community water systems at 180.

A4.4 Health effects assessment

In the dose–response assessment, the USEPA utilized the following:

- **Infection dose–response function:** The USEPA considered results from three human volunteer feeding studies – TAMU (collected from a veterinary student exposed during necropsy of an infected foal), Iowa (derived from a calf) and UCP (derived from a calf). A variant of the exponential dose–response model was fitted to the data sets using a Bayesian approach (mean value of $r = 0.036$).
- **Morbidity following infection:** For the purpose of this risk assessment, the morbidity rate was independent of dose, with a central tendency (mode) of 50%, a lower bound of 30% and an upper bound of 70%.
- **Mortality given that an illness has occurred:** Based on the cryptosporidiosis outbreak in Milwaukee, USA (50 deaths per 400 000 illnesses), this rate was adjusted to reflect changes in rates of illnesses and advanced treatments that have lessened mortality among persons living with HIV/AIDS. Further adjustments were used to reflect the differences between the populations living in areas served by filtered and unfiltered systems (a factor of 11.45 for unfiltered systems and a factor of 6.93 for filtered systems).

A4.5 Risk characterization

The mean annual cases of illness and death using each source water data set are summarized in Table A4.1. The differences between the pre-rule and post-rule cryptosporidiosis estimates were the basis for obtaining the “cases avoided”. The USEPA estimated that full implementation of the LT2ESWTR will reduce the incidence of cryptosporidiosis within the range of 84 609–464 069 mean illnesses per year (90% confidence bound across the three data sets ranging from 7 778 to 1 324 897 cases avoided; results not shown), with an associated reduction of 14–77 mean deaths (90% confidence bound across the three data sets of 1–219; results not shown). The additional *Cryptosporidium* treatment requirements of the LT2ESWTR will also reduce exposure to other microbial pathogens, such as *Giardia*, that co-occur with *Cryptosporidium*.

A4.6 Risk management

The QMRA provides a best estimate of the benefits expected in terms of cases of cryptosporidiosis avoided from the LT2ESWTR rule, providing a valuable risk communication tool. The assessment was also used to identify systems that would require corrective action and the nature of the actions likely to be implemented.

Table A4.1 Mean annual cases of illness and death estimated pre- and post-LT2ESWTR for each source water data set

	Estimated mean number of cases per year		
	Pre-LT2ESWTR	**Post-LT2ESWTR**	**No. of cases avoided**
Illnesses			
ICR data	491 091	27 022	464 069
ICRSSL data	147 185	62 575	84 609
ICRSSM data	257 985	59 559	198 426
Deaths			
ICR data	81	4	77
ICRSSL data	24	10	14
ICRSSM data	43	10	33

Source: adapted from USEPA (2005)

A4.7 Evaluation of the QMRA

Microbial risk assessment, comprising a two-dimensional Monte Carlo simulation model, was used to characterize the expected incidence of adverse health effects associated with exposure to *Cryptosporidium* and to quantify the benefit estimates of implementing the LT2ESWTR by calculating the difference between illness and death estimates for the baseline condition (pre-LT2ESWTR) and after implementation of the rule.

The difference between the results for data sets 1, 2 and 3 (Table A4.1) highlights the importance of the assumed *Cryptosporidium* source water concentration to the risk assessment results and hence the uncertainty associated with the selection of the most appropriate data set.

A4.8 Reference

- USEPA (2005). Economic analysis for the Final Long Term 2 Enhanced Surface Water Treatment Rule. Washington (DC): United States Environmental Protection Agency, Office of Water (http://nepis.epa.gov/Exe/ZyPDF.cgi/901S0000.PDF?Dockey=901S0000.PDF, accessed 19 April 2016).

A5 Case-study 5: Guidelines for water recycling – Setting health-based performance targets and safe use of wastewater in Australia (NWQMS, 2006)

A5.1 Introduction

In Australia, the beginning of the 21st century has been characterized by increased pressure on freshwater supplies in most large cities and in many regional areas. It has been identified that water traditionally seen as wastewater, such as sewage effluent and stormwater, should be considered as a water resource and should be used more widely, particularly in situations where water is not required to be of drinking-water quality. As part of the National Water Quality Management Strategy, guidelines were needed that would provide a nationally consistent approach to the treatment and recycling of sewage, greywater and stormwater. This case-study illustrates the use of QMRA to support guidelines on the setting of health-based performance targets for the safe use of recycled water.

A5.2 Problem formulation

The purpose of the QMRA was to provide national guidelines on the setting of health-based performance targets to ensure the safe use of recycled water. Following the guidance contained within the GWEG (WHO, 2006), for a given magnitude of contamination in the source water and a defined use, the required level of treatment was quantified in order to meet the health-based target of 1×10^{-6} DALY pppy.

The scope of the QMRA was defined by:

- **Hazard identification:** The guidelines are intended to address all enteric pathogens that may be present in sewage and greywater, and therefore three reference pathogens were selected, one to represent each of the three pathogen groups:
  - **Bacteria:** *Campylobacter* was selected, as it is the most common cause of bacterial gastroenteritis in Australia.
  - **Viruses:** A reference virus that incorporated a combination of rotavirus and adenovirus characteristics, using occurrence data for adenovirus and dose–response data for rotavirus.
  - **Protozoa:** *Cryptosporidium* was selected, as it has high infectivity, is resistant to chlorination and is one of the most important waterborne human pathogens in the world.
- **Exposure pathways:** The source material was sewage and/or greywater, treated and then used for a range of purposes, as illustrated in Fig. A5.1. Consideration was given to both intended and unintended uses.
- **Health outcomes:** After consideration of preventive measures, residual risk should be less than 10^{-6} DALY pppy.

A5.3 Exposure assessment

A5.3.1 Source

A deterministic QMRA was recommended, using point values to represent the inflow pathogen concentration. Given variability in the concentration of pathogens in wastewater and greywater, the upper 95th percentile was selected as the default point value:

- **Sewage concentration:** The 95th percentile of data from two Australian schemes (details of data sets not given) (unpublished data, South Australia Department of Health and Melbourne Water): *Campylobacter* = 7000 organisms·L^{-1}; adenovirus = 8000 organisms·L^{-1}; and *Cryptosporidium* = 2000 oocysts·L^{-1}.
- **Greywater concentration:** The arithmetic mean *E. coli* concentration was used as a surrogate for faecal contamination to determine the dilution of sewage in the greywater – that is, 10^5 *E. coli* per 100 mL (assumed to be equivalent to 1% sewage).

A5.3.2 Controls

To be defined within the QMRA in order to achieve health-based targets.

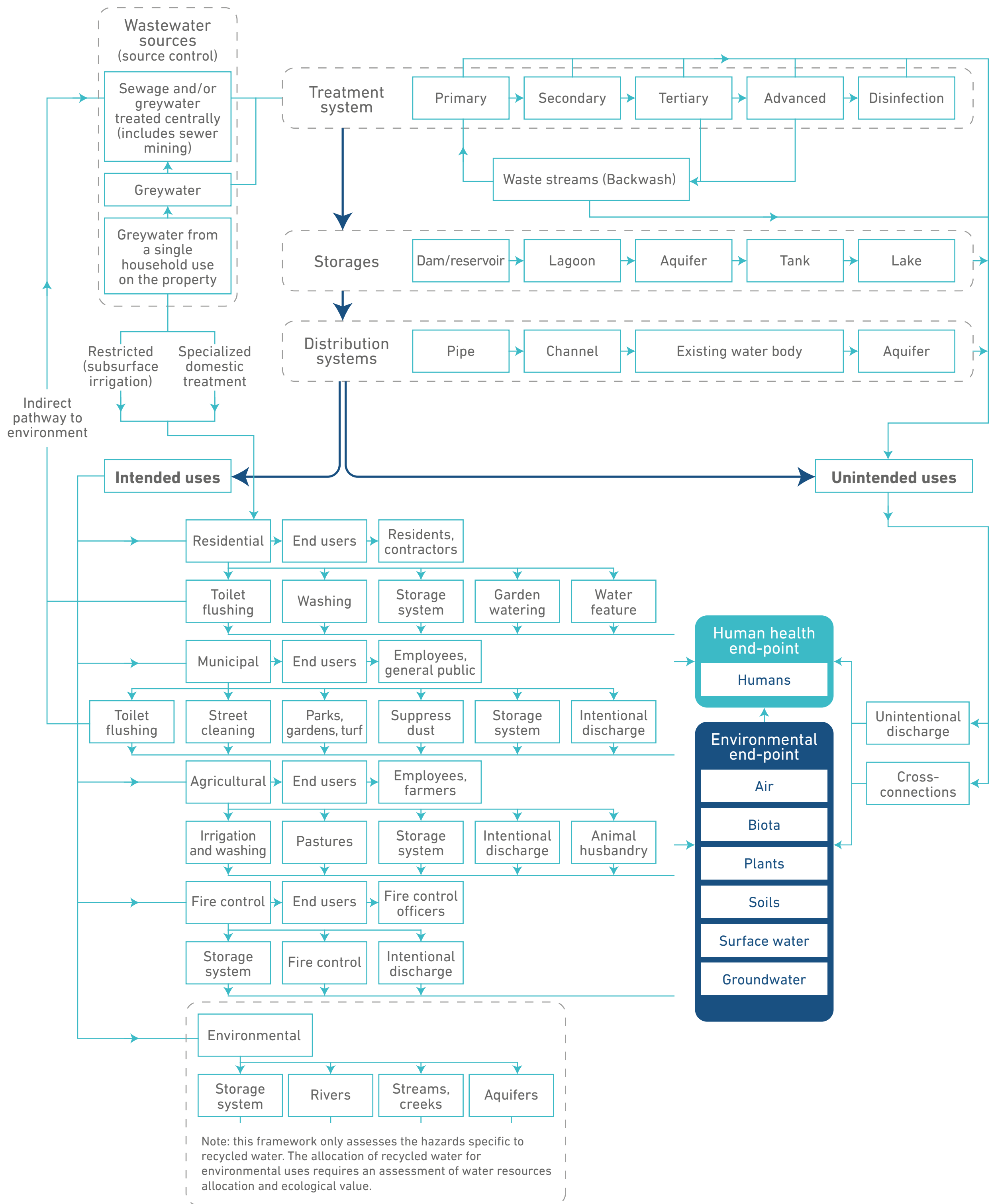


Fig. A5.1 Schematic illustration of water uses (including intended and unintended exposures) considered in the guidelines (adapted from NWQMS, 2006)

A5.3.3 Intake

Examples of intake volumes for the range of intended uses, based on a combination of scientific data and order of magnitude reference values, were compiled and provided in the guidelines (Table A5.1). These values were recommended as conservative defaults where specific or local information is not available.

Table A5.1 Intended uses and associated exposures for recycled water

Activity	Route of exposure	Volume (mL)	Frequency/ person/ year	Comments
Garden irrigation	Ingestion of sprays	0.1	90	Garden watering estimated to typically occur every second day during dry months (half year). Exposure to aerosols occurs during watering.
Garden irrigation	Routine ingestion	1	90	Routine exposure results from indirect ingestion via contact with plants, lawns, etc.
	Accidental ingestion	100	1	Infrequent event.
Municipal irrigation	Ingestion	1	50	Frequencies moderate, as most people use municipal areas sparingly (estimate 1/2–3 weeks). People are unlikely to be directly exposed to large amounts of spray, and therefore exposure is from indirect ingestion via contact with lawns, etc. Likely to be higher when used to irrigate facilities such as sports grounds and golf courses (estimate 1/week).
Food crop consumption (home grown)	Ingestion	5 (lettuce)	7	100 g of lettuce leaves hold 10.8 mL water and cucumbers 0.4 mL at worst case (immediately post-watering).[a] A serving of lettuce (40 g) might hold 5 mL of recycled water, and other produce might hold up to 1 mL per serving. Calculated frequencies are based on ABS data.[b]
		1 (other raw produce)	50	
Food crop consumption (commercial)	Ingestion	5 (lettuce)	70	100 g of lettuce leaves hold 10.8 mL water and cucumbers 0.4 mL at worst case (immediately post-watering).[a] A serving of lettuce (40 g) might hold 5 mL of recycled water, and other produce might hold up to 1 mL per serving. Calculated frequencies are based on ABS data.[c]
		1 (other raw produce)	140	
Toilet flushing	Ingestion of sprays	0.01	1 100	Frequency based on 3 uses of home toilet per day. Aerosol volumes are less than those produced by garden irrigation.
Washing machine use	Ingestion of sprays	0.01	100	Assumes one member of household exposed. Calculated frequency based on ABS data.[d] Aerosol volumes are less than those produced by garden irrigation (machines usually closed during operation).
Firefighting	Ingestion of water and sprays	20	50	Median ingestion for firefighters estimated at 20 mL per fire, with a maximum number of fires fought within area served by recycled water of 50 per year.[e]
Cross-connection of dual-reticulation systems with drinking-water mains	Ingestion	1 000/day	1/1 000 houses	Total consumption is assumed to be 2 L per day, of which 1 L is consumed cold.[f] Affected individuals may consume water 365 days per year. A conservative estimate of 1/1 000 houses has been considered.

ABS: Australian Bureau of Statistics

[a] Shuval, Lampert & Fattal (1997).

[b] ABS data show that 12% of households grow lettuce and 35% grow some type of produce (ABS, 1995); they also show that Australians eat leafy vegetables 140 times per year and eat other vegetables at a similar rate (ABS, 1995). Hence, it can be estimated that “other produce”, such as tomatoes, carrots, etc., in combination, are eaten 280 times per year. Watering with recycled water is used to augment rainfall. Assuming that watering occurs for 6 months of the year, frequency of consumption of lettuce irrigated with recycled water = 140 × 0.5 × 12%, and frequency of consumption of other raw produce = 280 × 0.5 × 35%.

[c] Using the same ABS data as in note b, frequency of consumption of lettuce irrigated with recycled water = 140 × 0.5 for lettuce, and frequency of consumption of other raw produce = 280 × 0.5.

[d] ABS data show average of 2.6 people per household (ABS, 2001). The amount of washing is estimated at five loads per week; therefore, the frequency = 5 × 52 ÷ 2.6

[e] Firefighting is an occupational exposure; the exposures were assessed by the Queensland Department of Emergency Services.

[f] WHO (2004).

Source: NWQMS (2006)

A5.4 Health effects assessment

Dose–response models recommended for use in defining the treatment requirements are included in Table A5.2, and the low-dose approximations are included in Table A5.3.

Table A5.2 Dose–response relationships for reference organisms

Organism type	Distribution	Model[a]	Parameters[b]
Enteric virus (rotavirus)	Beta-Poisson	$P_{inf} = 1-(1+d/\beta)^{-\alpha}$	$\alpha = 0.253$ $\beta = 0.426$
Bacterium (*Campylobacter jejuni*)	Beta-Poisson	$P_{inf} = 1-(1+d/\beta)^{-\alpha}$	$\alpha = 0.145$ $\beta = 7.58$
Protozoan (*Cryptosporidium parvum*)	Exponential	$P_{inf} = 1-\exp(-rd)$	$r = 0.059$

[a] α and r are parameters describing probability of infection; d = dose; β = median infective dose (N_{50}) ÷ ($2^{1/\alpha}-1$); P_{inf} = probability of infection.
[b] Model parameters are as described in Table 9.15 from Haas, Rose & Gerba (1999), except for *Cryptosporidium*, for which the data of Messner, Chappell & Okhuysen (2001) have been used.
Source: NWQMS (2006)

Table A5.3 Low-dose approximations for dose–response relationships for reference organisms

Organism type	Distribution	Model
Enteric virus (rotavirus)	Beta-Poisson low-dose approximation[a]	$P_{inf} = 0.59d$
Bacterium (*Campylobacter jejuni*)	Beta-Poisson low-dose approximation	$P_{inf} = 0.19d$
Protozoan (*Cryptosporidium parvum*)	Exponential low-dose approximation	$P_{inf} = 0.059d$

d: dose
[a] Low dose defined as less than 0.1 organism for *Cryptosporidium* and *Campylobacter* and less than 0.01 organism for rotavirus.
Source: NWQMS (2006)

A5.5 Risk characterization

The information from the exposure assessment and health effects assessment was combined to determine the magnitude of risk and to quantify the required level of treatment in order to meet the health-based target. The risk was therefore recommended to be assessed on two levels:

1) **Maximum risk:** risk in the absence of preventive measures;
2) **Residual risk:** risk that remains after consideration of existing preventive measures.

After consideration of preventive measures, residual risk should be less than 10^{-6} DALY pppy.

Performance targets for microbial hazards represent the reductions required to achieve a residual risk that complies with the tolerable level of 10^{-6} DALY pppy.

$\log_{10}$ reduction = $\log_{10}$ (concentration in source water × exposure × N ÷ DALYd) Eq. A5.1

where N is the frequency of exposure per year and DALYd is the dose equivalent to a DALY of 10^{-6} (1.6×10^{-2} *Cryptosporidium*, 2.5×10^{-3} rotavirus, 3.8×10^{-2} *Campylobacter*). DALYd includes consideration of dose–response and ratio of infection to illness.

Table A5.4 summarizes example calculations (full details of calculations included in Appendix 2 of NWQMS [2006]).

Table A5.4 shows that viruses require the highest $\log_{10}$ reductions. This reflects the high infectivity of viruses compared with bacteria and the higher disease burden of viruses compared with protozoa (in rare cases, rotavirus

infections can be fatal, and *Cryptosporidium* causes self-limiting diarrhoea with no long-term impacts for the general population). Table A5.4 also shows that the possibility of cross-connections represents a significant proportion of the exposure associated with dual-reticulation systems. Decreasing the likelihood of cross-connections would reduce the required log_{10} reductions.

Table A5.4 Log_{10} reductions for priority uses of recycled water from treated sewage

Activity	Route of exposure	Exposure (L) × frequency (per year)	Log_{10} reduction[a]		
			Crypto-sporidium	Rotavirus	*Campylo-bacter*
Commercial food crops	Ingestion – Lettuce Ingestion – Other produce	0.005 × 70 0.001 × 140			
	Total	**0.49**	**4.8**	**6.1**	**5.0**
Dual reticulation					
Garden irrigation	Ingestion of sprays Ingestion – Low Ingestion – High	0.000 1 × 90 0.001 × 90 0.1 × 1			
	Total	**0.2**	**4.4**	**5.8**	**4.6**
Garden food crops	Ingestion – Lettuce Ingestion – Other produce	0.005 × 7 0.001 × 50			
	Total	**0.09**	**4.0**	**5.3**	**4.2**
Internal uses					
Toilet flushing	Ingestion of sprays	0.000 01 × 1 100	3.1	4.5	3.3
Washing machine	Ingestion of sprays	0.000 01 × 100	2.1	3.5	2.3
Cross-connections	Ingestion	1 × 0.365	4.7	6.1	4.8
Total internal use (no garden use)		0.38	4.7	6.1	4.8
Total residential use (garden + internal)		0.67	4.9	6.3	5.1
Municipal irrigation	Ingestion of sprays	0.001 × 50	3.7	5.2	4.0
Dual reticulation plus municipal irrigation	Ingestion of water and sprays	0.72	5.0	6.4	5.1
Firefighting	Ingestion of water and sprays	0.02 × 50	5.1	6.5	5.3

[a] Log_{10} reduction calculations:

Cryptosporidium = Log_{10} (number of organisms in sewage × exposure (L) × frequency ÷ 1.6×10^{-2})

Rotavirus = Log_{10} (number of organisms in sewage × exposure (L) × frequency ÷ 2.5×10^{-3})

Campylobacter = Log_{10} (number of organisms in sewage × exposure (L) × frequency ÷ 3.8×10^{-2})

Source: NWQMS (2006)

A5.6 Risk management

Recycled water guidelines commonly specify combinations of treatment processes (Table A5.5) together with on-site controls (Table A5.6) and use restrictions to provide water of acceptable quality for identified uses. Using treatment as the primary means of minimizing risk from microbial hazards focuses control within a treatment plant. However, treatment is relatively expensive, and management of this type of facility requires a high degree of technical expertise.

Employing on-site controls and use restrictions reduces the focus on treatment. Controls can be used in combination with standard recycled water treatment processes that are often used for treating sewage (e.g. secondary treatment, storage lagoons and disinfection), with or without recycling of the final product. In this way, recycling can be introduced at existing facilities without the need for expensive retrofitting or treatment upgrades. However, when on-site controls and use restrictions are employed, preventive measures are spread over a much

broader area, and some measures might need to be implemented at a local user level. As a result, there is a greater need for observational monitoring, user education, surveillance and auditing.

The preventive measures chosen will be determined by issues such as:

- cost;
- intended use;
- existing treatment facilities;
- technical expertise;
- availability of land (e.g. if buffer zones are to be used);
- public access (e.g. use in tourist areas within capital cities compared with recycling in rural towns); and
- public perception and requirements.

Table A5.5 Indicative $\log_{10}$ removals of enteric pathogens and indicator organisms

Treatment	Indicative $\log_{10}$ reductions[a]							
	E. coli	Bacterial pathogens (including *Campylobacter*)	Viruses (including adenoviruses, rotaviruses and enteroviruses)	Phage	*Giardia*	*Cryptosporidium*	*Clostridium perfringens*	Helminths
Primary treatment	0–0.5	0–0.5	0–0.1	N/A	0.5–1.0	0–0.5	0–0.5	0–2.0
Secondary treatment	1.0–3.0	1.0–3.0	0.5–2.0	0.5–2.5	0.5–1.5	0.5–1.0	0.5–1.0	0–2.0
Dual-media filtration with coagulation	0–1.0	0–1.0	0.5–3.0	1.0–4.0	1.0–3.0	1.5–2.5	0–1.0	2.0–3.0
Membrane filtration	3.5–>6.0	3.5–>6.0	2.5–>6.0	3–>6.0	>6.0	>6.0	>6.0	>6.0
Reverse osmosis	>6.0	>6.0	>6.0	>6.0	>6.0	>6.0	>6.0	>6.0
Lagoon storage	1.0–5.0	1.0–5.0	1.0–4.0	1.0–4.0	3.0–4.0	1.0–3.5	N/A	1.5–>3.0
Chlorination	2.0–6.0	2.0–6.0	1.0–3.0	0–2.5	0.5–1.5	0–0.5	1.0–2.0	0–1.0
Ozonation	2.0–6.0	2.0–6.0	3.0–6.0	2.0–6.0	N/A	N/A	0–0.5	N/A
UV light	2.0–>4.0	2.0–>4.0	>1.0 adenovirus >3.0 enterovirus, hepatitis A virus	3.0–6.0	>3.0	>3.0	N/A	N/A
Wetlands – surface flow	1.5–2.5	1.0	N/A	1.5–2.0	0.5–1.5	0.5–1.0	1.5	0–2.0
Wetlands – subsurface flow	0.5–3.0	1.0–3.0	N/A	1.5–2.0	1.5–2.0	0.5–1.0	1.0–3.0	N/A

N/A: not available

[a] Reductions depend on specific features of the process, including detention times, pore size, filter depths, disinfectant.

Based on WHO (1989); Rose et al. (1996); NRC (1998); Bitton (1999); Rose et al. (2001); Mara & Horan (2003); USEPA (2003, 2004)

Source: NWQMS (2006)

A5.7 Evaluation of the QMRA

The approach adopted in the development of the Australian recycled water guidelines was necessarily conservative because of the many sources of uncertainty in the QMRA assumptions and expected $\log_{10}$ reductions achieved by different control measures. However, the QMRA provided a structured quantitative framework to allow for a diverse range of potential recycled water uses to be evaluated. The current level of scientific knowledge could be incorporated, with the opportunity for local refinement. Most importantly, the approach allowed for appropriate combinations of treatment options, and preventive measures could be selected to ensure that the water was fit-for-purpose.

Table A5.6 Exposure reductions provided by on-site preventive measures

Control measure	Reduction in exposure to pathogens
Cooking or processing of produce (e.g. cereal, wine grapes)	5–6 $\log_{10}$
Removal of skins from produce before consumption	2 $\log_{10}$
Drip irrigation of crops	2 $\log_{10}$
Drip irrigation of crops with limited to no ground contact (e.g. tomatoes, capsicums)	3 $\log_{10}$
Drip irrigation of raised crops with no ground contact (e.g. apples, apricots, grapes)	5 $\log_{10}$
Subsurface irrigation of aboveground crops	4 $\log_{10}$
Withholding periods – produce (decay rate)	0.5 $\log_{10}$/day[a]
Withholding periods for irrigation of parks/sports grounds (1–4 hours)	1 $\log_{10}$
Spray drift control (microsprinklers, anemometer systems, inward-throwing sprinklers, etc.)	1 $\log_{10}$
Drip irrigation of plants/shrubs	4 $\log_{10}$
Subsurface irrigation of plants/shrubs or grassed areas	5–6 $\log_{10}$
No public access during irrigation	2 $\log_{10}$
No public access during irrigation and limited contact after (non-grassed areas) (e.g. food crop irrigation)	3 $\log_{10}$
Buffer zones (25–30 m)	1 $\log_{10}$

[a] Based on virus inactivation. Enteric bacteria are probably inactivated at a similar rate. Protozoa will be inactivated if withholding periods involve desiccation.

Based on Asano et al. (1992); Tanaka et al. (1998); Haas, Rose & Gerba (1999); van Ginneken & Oron (2000); Petterson, Teunis & Ashbolt (2001); Mara & Horan (2003)

Source: NWQMS (2006)

A5.8 References

- ABS (1995). National Nutrition Survey: foods eaten. Canberra: Commonwealth of Australia, Australian Bureau of Statistics.
- ABS (2001). Census of population housing: selected social and housing characteristics. Canberra: Commonwealth of Australia, Australian Bureau of Statistics.
- Asano T, Leong LYC, Rigby MG, Sakaji RH (1992). Evaluation of the California wastewater reclamation criteria using enteric virus monitoring data. Water Sci Technol. 26(7–8):1513–24.
- Bitton G (1999). Wastewater microbiology. New York (NY): Wiley-Liss.
- Haas CN, Rose JB, Gerba CP (1999). Quantitative microbial risk assessment. New York (NY): John Wiley & Sons, Inc.
- Mara D, Horan N (2003). Handbook of water and wastewater microbiology. London: Academic Press.
- Messner MJ, Chappell CL, Okhuysen PC (2001). Risk assessment for *Cryptosporidium*: a hierarchical Bayesian analysis of human dose response data. Water Res. 35(16):3934–40.

- NRC (1998). Issues in potable reuse. Washington (DC): National Research Council, National Academy Press.
- NWQMS (2006). National Water Quality Management Strategy – National guidelines for water recycling: managing health and environmental risks. Report of the National Resource Management Ministerial Council, Environment Protection and Heritage Council and Australian Health Ministers' Conference (http://www.environment.gov.au/system/files/resources/044e7a7e-558a-4abf-b985-2e831d8f36d1/files/water-recycling-guidelines-health-environmental-21.pdf, accessed 9 April 2016).
- Petterson SR, Teunis PFM, Ashbolt N (2001). Modeling virus inactivation on salad crops using microbial count data. Risk Anal. 21(6):1097–1107.
- Rose JB, Dickson LJ, Farrah SR, Carnahan RP (1996). Removal of pathogenic and indicator microorganisms by a full-scale water reclamation facility. Water Res. 30(11):2785–97.
- Rose JB, Huffman DE, Riley K, Farrah SR, Lukasik JO, Hamann CL (2001). Reduction of enteric microorganisms at the upper Occoquan Sewage Authority water reclamation plant. Water Environ Res. 73(6):711–20.
- Shuval H, Lampert Y, Fattal B (1997). Development of a risk assessment approach for evaluating wastewater reuse standards for agriculture. Water Sci Technol. 35(11–12):15–20.
- Tanaka H, Asano T, Schroeder ED, Tchobanoglous G (1998). Estimating the safety of wastewater reclamation and reuse using enteric virus monitoring data. Water Environ Res. 70(1):39–51.
- USEPA (2003). LT1ESWTR disinfection profiling and benchmarking technical guidance manual. Washington (DC): United States Environmental Protection Agency, Office of Water.
- USEPA (2004). Guidelines for water reuse. Washington (DC): United States Environmental Protection Agency.
- van Ginneken M, Oron G (2000). Risk assessment of consuming agricultural products irrigated with reclaimed wastewater: an exposure model. Water Resour Res. 36:2691–9.
- WHO (1989). Health guidelines for the use of wastewater in agriculture and aquaculture. Report of a WHO Scientific Group. Geneva: World Health Organization (http://apps.who.int/iris/bitstream/10665/39401/1/WHO_TRS_778.pdf, accessed 20 April 2016).
- WHO (2004). Guidelines for drinking-water quality, third edition. Vol. 1. Geneva: World Health Organization.
- WHO (2006). Guidelines for the safe use of wastewater, excreta and greywater. Vol. 2. Wastewater use in agriculture. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/gsuweg2/en/, accessed 9 April 2016).

A6 Case-study 6: WHO health-based criteria for evaluating household water treatment technologies

A6.1 Introduction

Globally, an estimated 663 million people lack access to improved drinking-water sources (UNICEF/WHO, 2015); even where improved sources are available, the water supply is not necessarily safe. At least 1.9 billion people rely on either unimproved sources or improved sources that are faecally contaminated (WHO, 2014). It is in these settings that household water treatment (HWT) and safe storage (HWTS) can serve as an important interim measure to make drinking-water safer. Health gains from HWTS can be achieved only when treatment products are effective in removing pathogens from drinking-water and are used correctly and consistently. A variety of HWT products – with performance ranging from little to considerable pathogen removal – are available. In order to determine how well such products protect the health of users, WHO developed globally relevant health-based performance criteria for HWT using QMRA (WHO, 2011). The assumptions and methodology employed in developing the health-based performance targets for HWT are a direct application of the preventive risk-based framework outlined in the GDWQ (WHO, in preparation).

The functional objective of HWT is the removal of pathogens and other contaminants. Thus, QMRA, through the application of pathogen-specific dose–response information, provides a direct method for linking reductions in waterborne pathogens as a result of water treatment to health impacts.

A6.2 Problem formulation

The purpose of the QMRA was to inform the evaluation of HWT and to ensure that such technologies reduce waterborne pathogens sufficiently to protect health.

The scope of the QMRA was defined by:

- **Hazard identification:** The assessment was undertaken for the three main classes of pathogens (bacteria, viruses and protozoa) that cause diarrhoeal disease. For each class, a reference pathogen was selected, based on occurrence in human populations and faecally contaminated water, documented dose–response relationships and characterization in water. Thus, the three reference pathogens selected were the bacterium *Campylobacter jejuni*, rotavirus and the protozoan parasite *Cryptosporidium.* In addition, these three are cited in the GDWQ as key waterborne reference pathogens (WHO, in preparation).
- **Exposure pathways:** Contaminated untreated drinking-water; this could include surface water sources (pond or river), covered or uncovered communal wells and/or inadequately treated tap water.
- **Health outcome:** Annual probability of diarrhoeal illness.

A6.3 Exposure assessment

Although, ideally, local water quality data are used to inform QMRA, there is often a lack of reliable data in many countries where drinking-water is unsafe and unreliable and HWTS is used to improve water quality. Therefore, it was assumed that untreated, uncharacterized water is 0.01% wastewater. The concentrations of *Campylobacter jejuni*, rotavirus and *Cryptosporidium* in wastewater were derived from published estimates regarding the numbers of such pathogens excreted daily per infected person in a tropical community in a developing country (WHO, 2011; see Table A6.1). Finally, it was assumed that each person consumes 1 L of drinking-water each day (WHO, 2011; see Table A6.1). This resulted in the following daily exposure estimates: 1 (*Campylobacter jejuni*), 0.1 (*Cryptosporidium*) and 1 (rotavirus) (WHO, 2011; see Table A6.1).

A6.4 Health effects assessment

The exponential dose–response models cited in the GDWQ (WHO, in preparation) were used to conduct the health effects assessment. The probability of infection per organism was r = 0.019 for *Campylobacter jejuni*, r = 0.20 for *Cryptosporidium* and r = 0.59 for rotavirus (WHO, in preparation; see Table A6.1). One limitation of these models is that they were developed using a population from developed countries where immunity may differ from that of the populations using HWTS, which largely reside in developing countries.

Table A6.1 Calculation of required $\log_{10}$ reduction of microbes to achieve the "highly protective" performance level

	Units	*Cryptosporidium*	*Campylobacter jejuni*	Rotavirus
Raw water quality (C_R), assumed	Organisms per litre	0.1	1	1
Treatment efficacy required to reach tolerable risk (PT)	$\log_{10}$ reduction required	3.88	3.98	4.96
Drinking-water quality (C_D)	Organisms per litre	1.32×10^{-5}	1.05×10^{-4}	1.10×10^{-5}
Consumption of drinking-water (V)	Litres per person per day	1	1	1
Exposure by drinking-water (E)	Organisms per day ingested	1.34×10^{-5}	1.04×10^{-4}	1.10×10^{-5}
Dose–response (r)	Probability of infection per organism	0.20	0.019	0.59
Risk of infection ($P_{inf,d}$)	Per day	2.67×10^{-6}	1.99×10^{-6}	6.53×10^{-6}
Risk of infection ($P_{inf,y}$)	Per year	9.74×10^{-4}	7.25×10^{-4}	2.38×10^{-3}
Risk of diarrhoeal illness given infection ($P_{ill\|inf}$)		0.7	0.3	0.5
Risk of diarrhoeal illness (P_{ill})	Per year	6.82×10^{-4}	2.18×10^{-4}	1.19×10^{-3}
Disease burden (db)	DALYs per case	1.47×10^{-3}	4.60×10^{-3}	1.40×10^{-2}
Susceptible fraction (f_s)	Percentage of population	100%	100%	6%
Health outcome target	DALYs per year	1×10^{-6}	1×10^{-6}	1×10^{-6}
Formulae	$C_D = C_R \div 10^{PT}$ $E = C_D \times V$ $P_{inf,d} = E \times r$ $P_{ill} = P_{inf,y} \times P_{ill\|inf}$ $DB = P_{ill} \times db \times f_s \div 100$			

Source: adapted from WHO (in preparation). The format and calculations contained in this table follow the same approach as described in the GDWQ.

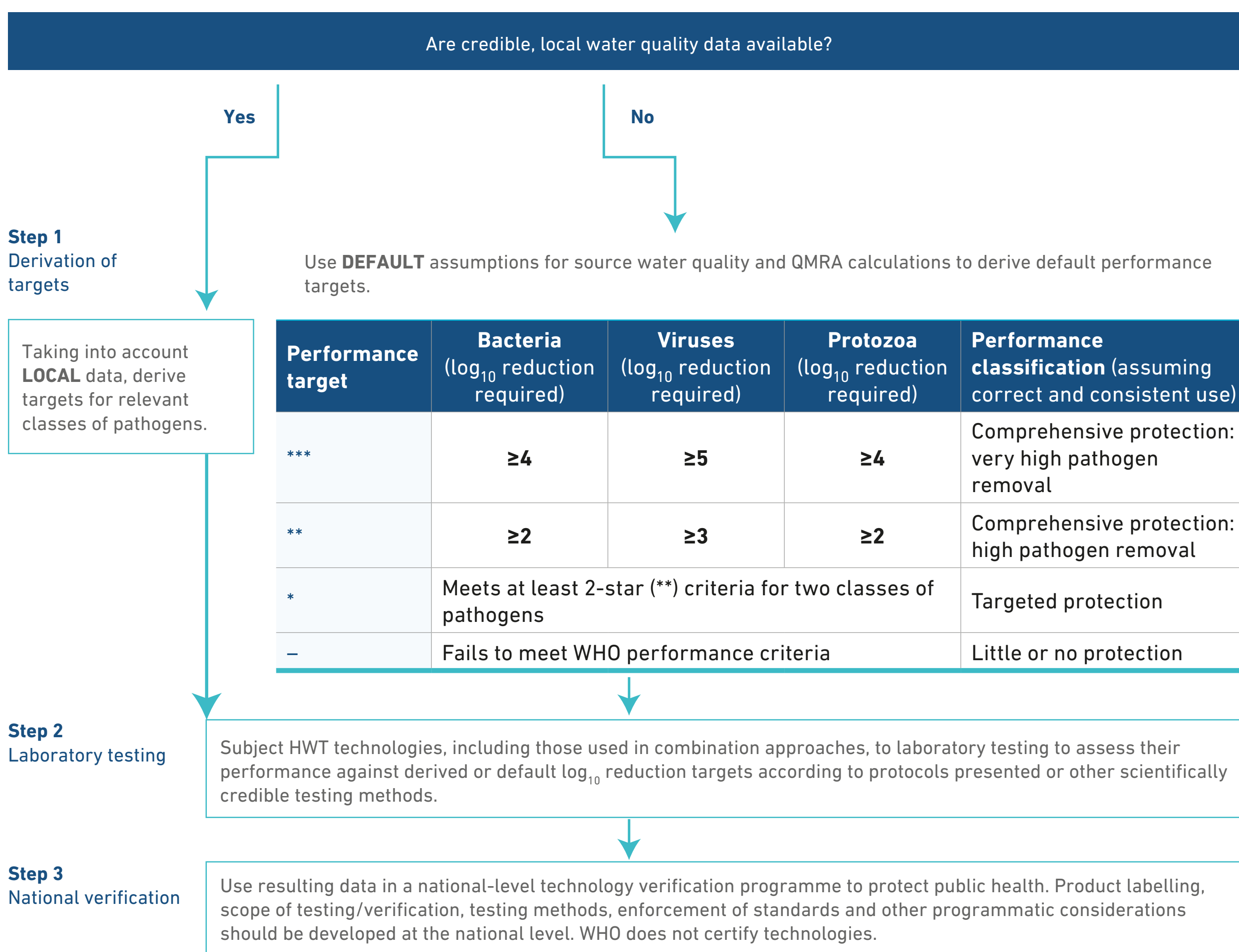
A6.5 Risk characterization

Treatment targets were defined for each pathogen class for three different levels of performance. The top tier, designated by three stars, indicates very high removal of pathogens, which represents those technologies that, if used correctly and consistently over an entire year, will limit the drinking-water diarrhoeal disease burden to 10^{-6} DALY pppy. The second tier, designated by two stars, defines pathogen removals that achieve a health-based target of 10^{-4} DALY pppy. Both three-star and two-star performances provide comprehensive protection against the three main classes of pathogens. Technologies in these two categories provide the greatest protection, especially in settings where:

- the burden of diarrhoeal disease is high owing to several classes of pathogens;
- there is no information on the specific pathogens in drinking-water, and a prudent approach is to protect against all three classes; and
- there is a lack of safe faecal management and safe piped water supplies.

The one-star classification clearly provides less protection; in general, these products should be used in targeted situations where there is a pathogen of known concern (e.g. in a cholera outbreak) or in combination with other products to ensure comprehensive protection.

In order to achieve these performance levels, technologies may be used in combination. For example, using a ceramic filter followed by chlorination provides a multibarrier approach that achieves a greater level of performance than either technology could achieve alone.



Performance target	Bacteria (log_{10} reduction required)	Viruses (log_{10} reduction required)	Protozoa (log_{10} reduction required)	Performance classification (assuming correct and consistent use)
***	≥4	≥5	≥4	Comprehensive protection: very high pathogen removal
**	≥2	≥3	≥2	Comprehensive protection: high pathogen removal
*	Meets at least 2-star (**) criteria for two classes of pathogens			Targeted protection
–	Fails to meet WHO performance criteria			Little or no protection

Fig. A6.1 Process for achieving health-based performance targets (adapted from WHO, 2011, 2016c)

A6.6 Risk management

Fig. A6.1 illustrates the overall process for establishing health-based performance targets in regards to HWT. Each step is part of a sequential process to identify performance targets (where the performance level of 10^{-6} (***) reflects the log_{10} reduction targets included in the GDWQ, as shown in Table A6.1), conduct testing and, finally, inform national verification efforts. Step 1, as described, proposes three levels of targets, unless sufficient local data are available to derive such targets. However, in most countries where drinking-water quality is compromised and HWT plays a potentially important role, there are neither sufficient data nor the laboratory capacity and expertise to generate such data.

The actual testing of technologies against the targets is Step 2. Recognizing that many countries have neither the resources nor the capacity to evaluate HWT performance according to WHO performance criteria, WHO has established the International Scheme to Evaluate Household Water Treatment Technologies (WHO, 2016a). WHO works with an Independent Advisory Committee of experts and designated testing laboratories to consistently and independently evaluate the performance of HWT products and, in so doing, guide Member States, procuring agencies and users in product selection. Further, WHO has developed an overall harmonized test protocol as well as specific test protocols for the main types of technologies (WHO, 2016b). The protocols employ both general test water and challenge test water, the latter representing highly turbid surface water sources. These protocols are used by the WHO Scheme's designated testing laboratories and are being adapted for use in low-income settings, such as Ethiopia, where laboratory facilities and expertise are limited. Increasingly, WHO, with advice from experts on the Independent Advisory Committee, is working to identify surrogates for each pathogen class rather than actual pathogens to assess performance. This is important not only for reducing the cost and time of testing, but also to allow for a wider range of laboratories and countries to conduct testing.

Finally, in Step 3, testing should be linked to and inform national regulation and technology verification efforts. This is a critical step for ensuring that both governments and consumers are able to make informed choices about the protective effect of a range of technologies. WHO is supporting a number of countries to strengthen national regulations and evaluation in order to ensure that testing results are appropriately considered when selecting HWT options. In addition, through the WHO/United Nations Children's Fund (UNICEF) International Network on Household Water Treatment and Safe Storage, WHO is working with other stakeholders to better understand and address the enabling environment to support smarter HWT implementation and improved health.

Further information on the WHO Scheme, including a full list of products tested or being tested and testing results, can be found on the WHO Scheme website: http://www.who.int/water_sanitation_health/water-quality/household/scheme-household-water-treatment/en/.

Further information on the UNICEF/WHO Household Water Treatment Network can be found at: http://www.who.int/water_sanitation_health/water-quality/household/household-water-network/en/.

A6.7 Evaluation of the QMRA

The principal objective of the QMRA was to derive health-based performance criteria for HWT systems at different tier levels. To be able to derive a generic performance target, it was necessary to make an assumption about the type and appropriate concentration of reference pathogens in source waters for HWT systems.

One critical consideration in utilizing QMRA to develop health-based performance targets is that the predicted health outcomes depend on the correct and consistent use of HWT, resulting in safe drinking-water. This puts the onus on the user, as opposed to centralized drinking-water treatment systems, where, assuming adequate treatment and management, there is less chance of error/non-adherence. Correct and consistent use is a challenge and one that is somewhat opaque, in part because many of the data do not directly measure use of HWT and because of the interconnected behavioural and economic factors that influence the use of HWT. However, monitoring is improving, with implementers increasingly assessing water quality or using remote sensors to determine correct and consistent use of HWT, for example, filters (WHO/UNICEF, 2012). In studies in which consistent use (adherence) was measured, it was often less than 70% (Arnold & Colford, 2007). According to a recent study that applied QMRA in the context of HWT, unless adherence is 90% or greater, tangible health gains are unlikely to be realized (Brown & Clasen, 2012). This highlights the importance of creating an enabling environment that supports high adherence. Furthermore, it also indicates that those technologies with a high user burden, with recurrent costs or that involve substantial behavioural change are unlikely to achieve health gains, even if they are microbiologically highly efficacious. Meanwhile, implementers are increasingly targeting HWTS towards those populations most at risk, such as pregnant mothers and those living with HIV. These populations are particularly aware of the dangers of consuming unsafe water and are more receptive to messages about the importance of safe drinking-water. In such populations, use is higher than 90% (Peletz et al., 2012; Woods, Foster & Kols, 2012). In addition, implementers are increasingly using evidence-based behavioural change approaches as an effective means to increase and sustain use of HWTS (Mosler, 2012).

A6.8 References

- Arnold BF, Colford JM Jr (2007). Treating water with chlorine at point-of-use to improve water quality and reduce child diarrhea in developing countries: a systematic review and meta-analysis. Am J Trop Med Hyg. 76:354–64.
- Brown J, Clasen T (2012). High adherence is necessary to realize health gains from water quality interventions. PLoS One. 7(5):e36735.
- Mosler HJ (2012). A systematic approach to behavior change interventions for the water and sanitation sector in developing countries: a conceptual model, a review, and a guideline. Int J Environ Health Res. 22(5):431–49.
- Peletz R, Simunyama M, Sarenje K, Baisley K, Filteau S, Kelly P et al. (2012). Assessing water filtration and safe storage in households with young children of HIV-positive mothers: a randomized, controlled trial in Zambia. PLoS One. 7(10):e46548.
- UNICEF/WHO (2015). Progress on sanitation and drinking water: 2015 update and MDG assessment. Geneva: World Health Organization and United Nations Children's Fund (UNICEF) (http://www.who.int/water_sanitation_health/publications/jmp-2015-update/en/, accessed 9 April 2016).
- WHO (2011). Evaluating household water treatment options: health-based targets and microbiological performance specifications. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/household_water/en/, accessed 9 April 2016).

- WHO (2014). Preventing diarrhoea through better water, sanitation and hygiene. Exposures and impacts in low- and middle-income countries. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/preventing-diarrhoea/en/, accessed 9 April 2016).
- WHO (2016a). WHO International Scheme to Evaluate Household Water Treatment Technologies [website]. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/water-quality/household/scheme-household-water-treatment/en/, accessed 4 May 2016).
- WHO (2016b). Scheme harmonized testing protocol. Geneva: World Health Organization (http://www.who.int/household_water/scheme/HarmonizedTestProtocol.pdf?ua=1, accessed 30 September 2015).
- WHO (2016c). Results of Round I of the WHO International Scheme to Evaluate Household Water Treatment Technologies. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/household-water-treatment-report-round-1/en/, accessed 4 May 2016).
- WHO (in preparation). Guidelines for drinking-water quality, fourth edition incorporating first addendum. Geneva: World Health Organization.
- WHO/UNICEF (2012). Toolkit for monitoring and evaluating household water treatment and safe storage. Geneva: World Health Organization/United Nations Children's Fund (http://www.who.int/water_sanitation_health/publications/toolkit_monitoring_evaluating/en/, accessed 30 September 2015).
- Woods S, Foster J, Kols A (2012). Understanding why women adopt and sustain home water treatment: insights from the Malawi antenatal care program. Soc Sci Med. 75:634–42.

B1 Hypothetical case-study

The following example is intended to demonstrate the consideration of variability and uncertainty within a QMRA of a drinking-water system. A single system is assessed several times using the same framework; however, the number of data and quantitative effort applied in the variability and uncertainty analysis for the exposure assessment are incrementally increased. It is intended as an illustrative hypothetical case-study only and is not intended to be prescriptive.

B1.1 Case description

> *Water is drawn from a large river and is treated by conventional treatment (coagulation/flocculation/sedimentation/rapid sand filtration) followed by disinfection with free chlorine at 1 mg·L^{-1}. Water is distributed to the consumer by a gravity-fed pressurized pipe network. The regulator has stipulated that the annual DALYs to the community from* Cryptosporidium *via drinking-water must not exceed 1 × 10^{-6} per person per year.*

With respect to *Cryptosporidium*, is the water supply chain able to deliver water of a quality that meets the regulator's health-based target?

The application of the QMRA framework to the case-study scenario is illustrated in Fig. B.1.

B1.2 Problem formulation

The purpose of the risk assessment was to evaluate whether or not the risk from *Cryptosporidium* is greater than or less than the health-based target of 1 × 10^{-6} DALY pppy. The scope of the risk assessment is defined by the hazard identification, the exposure pathways and the health outcomes:

- **Hazard identification:** The reference pathogen as defined by the regulator is *Cryptosporidium*.
- **Exposure pathways:** The exposure pathway is illustrated in Fig. B.1. Water is drawn from a river draining an impacted catchment (unknown magnitude) with many sources of *Cryptosporidium* (most likely human and

animal in origin). Water is treated by two engineered treatment barriers: conventional treatment (which could be considered as separate subprocesses of coagulation, flocculation, sedimentation and rapid sand filtration) and free chlorine disinfection. A third barrier, and potential site of recontamination, is the distribution system. Consumers are potentially exposed to *Cryptosporidium* via the consumption of unboiled tap water (boiling of tap water is assumed to inactivate *Cryptosporidium*).

- **Health outcomes:** The health outcome stipulated by the regulator is the annual DALY.

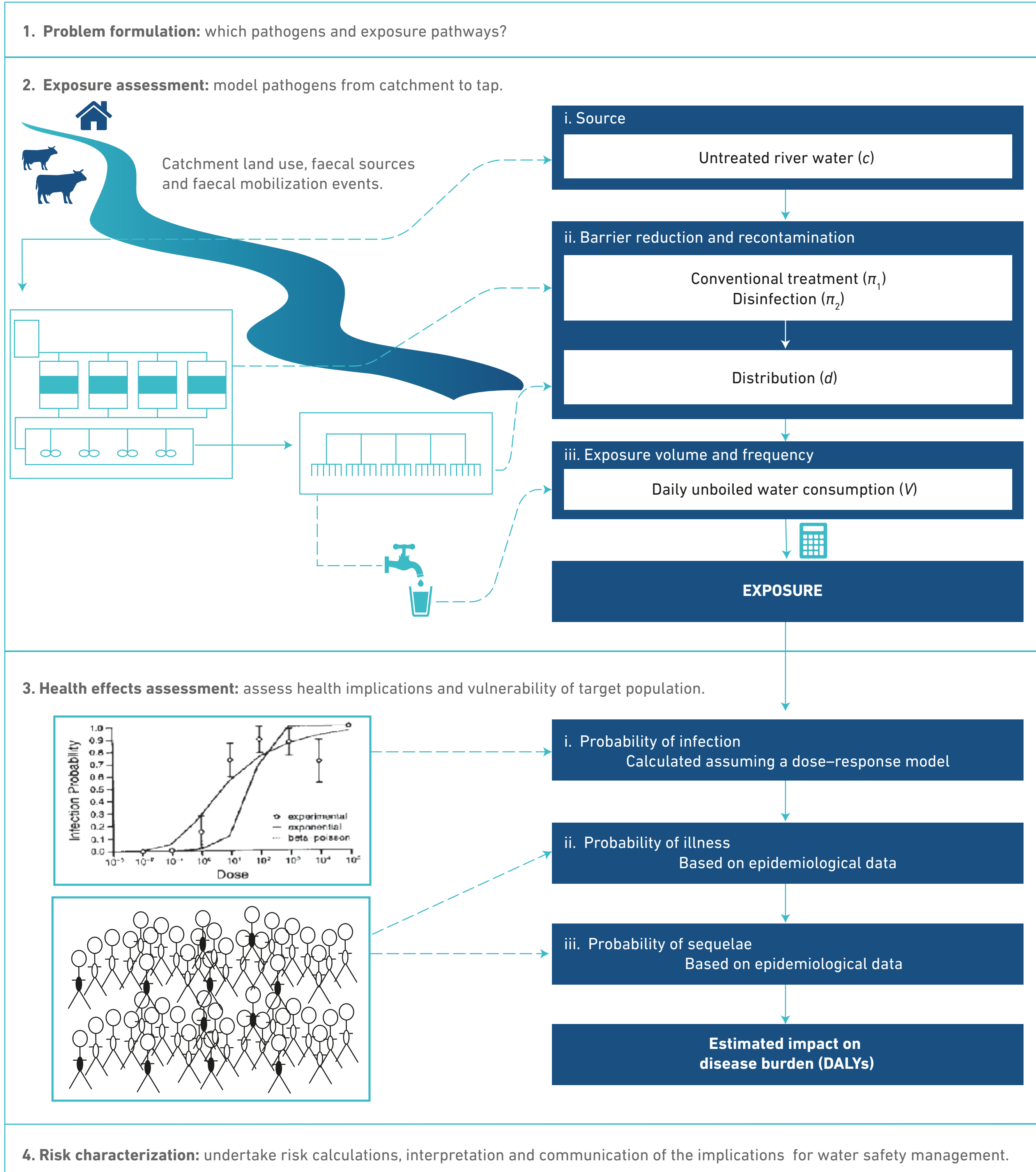


Fig. B.1 Schematic illustration of the QMRA framework applied to the drinking-water system in the hypothetical case-study

B1.3 Exposure assessment

The exposure pathway is illustrated in Fig. B.1. Each component of the exposure assessment (c, π_1, π_2, d and V) needs to be quantified based on available data. The removal is represented as π (the probability of individual pathogen passage, which is related to the $\log_{10}$ reduction [LR] of the pathogen population by: $\pi = 10^{-LR}$). In each part of this worked example, a separate exposure assessment will be undertaken using the same basic framework, but differing levels of detail in quantifying variability and uncertainty.

B1.4 Health effects assessment

The health outcome required by the regulator was the annualized DALY. Two pathogen-specific factors are required for the risk assessment: the dose–response relationship and the DALY per infection. Based on a review of the available models, the exact Beta-Poisson model fitted to the combined isolate data (see section 7.1 and Annex D; Teunis, Chappell & Okhuysen, 2002) was selected. The exact Beta-Poisson model was approximated at low doses using the exponential model:

$$P_{inf} = 1 - e^{r \cdot D} \qquad \text{Eq. B.1}$$

with: $r = \dfrac{\alpha}{\alpha + \beta} = \dfrac{0.115}{0.115 + 0.176} = 0.4$; and D = exposure dose (oocysts)

The DALY weighting used in the GDWQ (WHO, in preparation) was selected. This health effects assessment is consistent for all parts of the worked example, with the exception of the uncertainty analysis in section B6.

B1.5 Risk characterization

The exposure assessment and the health effects assessment were combined to calculate the probability of infection for the study system. The annualized DALY benchmark (1×10^{-6} DALY pppy) was transformed using the number of days of exposure to obtain the acceptable average daily infection probability. Given the assumptions of the GDWQ (WHO, in preparation), the average daily probability of infection is 2.6×10^{-6}. The risk was characterized separately for each part of the following worked example, depending on the assumptions of the exposure
assessment.

B2 Part A: Quantifying exposure with no local information – point estimates

The first-tier assessment was undertaken based on limited local information and predictions made using literature data without any local monitoring data.

B2.1 Data collection

- **Source:** The first-tier risk assessment was undertaken with only a description of catchment land use: a large impacted river with many faecal sources, including large urbanized areas, intensive agriculture and treated wastewater discharges. Catchment management practices for agricultural activities were unknown. Drawing on a review of information from the literature (Medema et al., 2009), only a vague prediction of *Cryptosporidium* concentration can be made without a thorough sanitary survey and/or microbial data. The river could reasonably be classified between the categories of "moderately polluted" (0.1 oocyst·L^{-1}) and "heavily polluted" (10 oocysts·L^{-1}).
- **Treatment:** The drinking-water treatment plant is operated by a water supply company that is subject to tight regulatory requirements. Turbidity in finished water prior to chlorination may not exceed 0.2 NTU in the combined effluent, and historical online turbidity measurements indicate that this requirement is achieved more than 95% of the time. No additional information on plant performance was available. A meta-analysis of published data on conventional treatment processes, undertaken by Hijnen & Medema (2007), reported a mean elimination capacity (MEC) for conventional treatment of 3.2 $\log_{10}$.
- **Distribution:** Water is distributed to the consumer by a gravity-fed pressurized pipe network. The system is well maintained; when leakages/breakages are identified, repairs are undertaken in accordance with strict quality assurance procedures that have been developed for protection of the water quality. No additional data on the distribution system were available.
- **Exposure:** No local data on the consumption of unboiled tap water were available.

B2.2 Quantification of model inputs

- **Source:** In the absence of any specific local data, a conservative upper bound on the *Cryptosporidium* concentration (10 oocysts·L^{-1}) was selected.
- **Treatment:** While there were limited data available on the performance of the treatment plant, it was concluded that it was reasonable to assume that the conventional treatment process was well operated. The MEC from reviewed studies was selected to describe the performance of conventional treatment. Free chlorine disinfection was assumed to be ineffective against *Cryptosporidium*, and therefore the probability of passage (see section 6.2.2) for the disinfection step was set equal to 1.
- **Distribution:** The impact of the distribution system was unknown. For the initial assessment, the distribution system was ignored.
- **Exposure:** In the absence of any local data, a reference value for exposure volume was selected at 1 L per person per day.

The inputs for the exposure assessment of Part A (section B2) are summarized in Table B.1.

Table B.1 Input values for exposure assessment: Part A

Component of exposure pathway[a]	Variable	Information source	Input value
i	*c*	Literature reports an average concentration up to 10 oocysts·L^{-1} for heavily polluted rivers	10
ii	π_1[b]	Meta-analysis of published studies on conventional treatment MEC = 3.2 $\log_{10}$ ($10^{-3.2}$ = 0.000 63)	0.000 63
	π_2	Free chlorine is ineffective against oocysts	1
	d	No available information	Effect ignored
iii	*V*	No available information – reference value selected (L)	1
	Exposure dose (*D*): $C_{source} \times (\pi_1) \times (\pi_2) \times V = 6.31 \times 10^{-3}$		

[a] See Fig. B.1.

[b] π represents probability of passage.

B2.3 Risk characterization

The daily probability of infection was calculated as:

$$P_{inf} = 1 - e^{-0.4 \times 6.31 \times 10^{-3}}$$

$$P_{inf} = 1.46 \times 10^{-3}$$

The calculated daily probability of infection is well above the daily average target of 2.6×10^{-6}. The calculated risk is very high, around 3 $\log_{10}$ higher than the target, and it is likely that the process train is inadequate to meet the target. An additional barrier or barriers may be required.

Were the input assumptions reasonable? The risk assessment was not based on any local water quality data and used very conservative initial oocyst concentration assumptions from the literature. Further analysis with local water quality data may be worthwhile before concluding that the water is unsafe and investing in additional treatment.

B3 Part B: Quantifying exposure with some local information – point estimates

B3.1 Data collection

Samples (n = 10) were collected to evaluate the concentration of *Cryptosporidium* in the river water. The average concentration of the 10 samples was reported to be 0.2 oocyst·L^{-1}. In addition, a historical data set of *E. coli*

concentrations collected over several years at the same location (n = 421) was obtained; the average *E. coli* concentration was 54 MPN·100 mL^{-1}. Comparison with literature data (Medema et al., 2009) indicates that these results suggest that the water source was more likely to be in the lower range of contamination identified in section B2: "moderately polluted" (0.1 oocyst·L^{-1}).

B3.2 Quantification of model inputs

The microbiological water quality data support the selection of a lower *Cryptosporidium* concentration for the source water, and a value of 0.2 oocyst·L^{-1} was implemented in the QMRA model. All other components of the model remained unchanged from section B2.

The inputs for the exposure assessment of Part B (section B3) are summarized in Table B.2.

Table B.2 Input values for exposure assessment: Part B

Component of exposure pathway[a]	Variable	Information source	Input value
i	c	Reported average concentration from local sampling (n = 10)	0.2
ii	π_1	Meta-analysis of published studies on conventional treatment MEC = 3.2 $\log_{10}$ ($10^{-3.2}$ = 0.000 63)	0.000 63
	π_2	Free chlorine is ineffective against oocysts	1
	d	No available information	Effect ignored
iii	V	No available information – reference value selected (L)	1
	Exposure dose (D): $C_{source} \times (\pi_1) \times (\pi_2) \times V = 1.26 \times 10^{-4}$		

[a] See Fig. B.1.

B3.3 Risk characterization

The daily probability of infection is calculated as:

$$P_{inf} = 1 - e^{-0.4 \times 1.26 \times 10^{-4}}$$

$$P_{inf} = 4.99 \times 10^{-5}$$

Recalculation of the P_{inf} with the new concentration of 0.2 oocyst·L^{-1} resulted in a much lower probability of infection, indeed 2 $\log_{10}$ lower; however, it is still approximately 1 $\log_{10}$ above the target of 2.6×10^{-6}. When relying on literature alone without site-specific data, it is often necessary to be conservative, and therefore high numbers need to be selected. A great deal of benefit can be obtained with local data to verify how realistic those initial assumptions are. In this case, the oocyst counts combined with the larger *E. coli* data set provided evidence that the expected level of faecal contamination was considerably lower than first thought. While the infection probability is still above the target, it is much closer and within an order of magnitude. The calculated value is the average or expected risk; however, this will vary between days. Investigation into the nature of that variation may assist with deciphering what is driving the risk associated with this system and what options may be available to reduce the risk below the target.

B4 Part C: Accounting for variability in exposure – point estimates

As the average risk was above the target, it was deemed worthwhile to tease out the influences of variability on the risk calculations. In this example, the variability was investigated using point estimates to describe different discrete conditions that the system may experience.

B4.1 Variation in source water concentration

B4.1.1 Data collection

Rather than rely on the average oocyst concentration, the 10 individual reported concentrations were obtained to investigate the variability in concentration. While on most occasions no oocysts were found (<0.1 oocyst·L^{-1}), on one occasion, a concentration of 1.9 oocysts·L^{-1} was reported.

B4.1.2 Quantification of model inputs

The concentration of oocysts was characterized under two conditions: nominal (background) conditions, where the concentration was assumed to be equal to half the detection limit (0.05 oocyst·L^{-1}), and peak conditions, where the concentration of 1.9 oocysts·L^{-1} was applied.

Input values for the exposure assessment are summarized in Table B.3.

Table B.3 Input values for exposure assessment: Part C (Variation in source water concentration)

Component of exposure pathway[a]	Variable	Information source	Input value	
			Nominal	Peak
i	*c*	Reported concentration under background and peak conditions from local sampling	0.05	1.9
ii	π_1	Meta-analysis of published studies on conventional treatment MEC = 3.2 $\log_{10}$ ($10^{-3.2}$ = 0.000 63)	0.000 63	
	π_2	Free chlorine is ineffective against oocysts	1	
	d	No available information	Effect ignored	
iii	*V*	No available information – reference value selected (L)	1	
	Exposure dose (*D* nominal): $C_{source} \times (\pi_1) \times (\pi_2) \times V = 3.15 \times 10^{-5}$ Exposure dose (*D* peak): $C_{source} \times (\pi_1) \times (\pi_2) \times V = 1.20 \times 10^{-3}$			

[a] See Fig. B.1.

B4.1.3 Risk characterization

The daily probability of infection is calculated as:

P_{inf} nominal $= 1 - e^{-0.4 \times 3.15 \times 10^{-5}} = 1.25 \times 10^{-5}$

P_{inf} peak $= 1 - e^{-0.4 \times 1.20 \times 10^{-8}} = 4.74 \times 10^{-4}$

Based on the range of measured *Cryptosporidium* concentrations, the risk varies greatly between days and is often below the detection limit, which, under the current assumptions of the model, is approximately the value of the allowable daily probability of infection. It may be that the overall risk to consumers could be most effectively reduced by managing the peak source water events. However, the risk assessment has not yet considered the variability in treatment.

B4.2 Variation in source water concentration, treatment and consumption volume

B4.2.1 Data collection

- **Treatment:** The water utility collected data at the treatment plant on aerobic spore removal across the conventional treatment barrier and reported that removal varied from 1.1 $\log_{10}$ to 4 $\log_{10}$ (mean = 2.5 $\log_{10}$). In the risk assessment, it was assumed that this reflected the (variation in) removal of *Cryptosporidium* oocysts by conventional treatment.

- **Exposure:** A review of a local drinking-water consumption study for the exposed population demonstrated that while some individuals consumed up to 1 L per day, the average consumption for the population was 0.15 L per day.

B4.2.2 Quantification of model inputs

The variability in both the treatment and exposure volume was quantified using point values to represent nominal and peak conditions. For treatment, nominal conditions were assumed equal to the mean $\log_{10}$ reduction based on the aerobic spore data, with the peak risk removal (i.e. poor removal) equal to the minimum reported removal = 1.1 $\log_{10}$. For exposure volumes, the nominal exposure volume was reduced to represent the average from the consumption study, and the 1 L per day was included as the peak (maximum) consumption.

Input values for the exposure assessment are summarized in Table B.4.

Table B.4 Input values for exposure assessment: Part C (Variation in source water concentration, treatment and consumption volume)

Component of exposure pathway[a]	Variable	Information source	Input value	
			Nominal	Peak
i	c	Reported concentration under background and peak conditions from local sampling	0.05	1.9
ii	π_1	Local data on aerobic spore removal range from 1.1 $\log_{10}$ to 4 $\log_{10}$ (mean 2.5 $\log_{10}$)	0.003 2 (2.5 $\log_{10}$)	0.079 (1.1 $\log_{10}$)
	π_2	Free chlorine is ineffective against oocysts	1	
	d	No available information	Effect ignored	
iii	V	Local drinking-water consumption study, average = 0.15 (L)	0.15	1

[a] See Fig. B.1.

B4.2.3 Risk characterization

When undertaking the calculations, the nature of the variability (between individuals or over time) needs to be accounted for. Source water and treatment performance inputs are variable over time; only one inflow concentration and treatment removal are assumed to occur. Conversely, variation in consumption volume refers to variability between individuals across the entire population. For a given quality of water delivered to the tap, some will consume a small portion of that water, and some will consume more. The combinations of conditions that are considered in the risk calculations, with the associated risk quantification, are included in Table B.5.

Table B.5 Quantifying risk under a range of different conditions: Part C (Variation in source water concentration, treatment and consumption volume)

Model input	Condition						
	Peak inflow concentration		Suboptimal conventional treatment performance		Combined peak and suboptimal performance		Baseline
i) Source water	Peak		Nominal		Peak		Nominal
ii) Conventional treatment	Nominal		Peak		Peak		Nominal
iii) Consumption volume	Nominal	Peak	Nominal	Peak	Nominal	Peak	Nominal
P_{inf}	3.6×10^{-4}	2.4×10^{-3}	2.4×10^{-4}	1.6×10^{-3}	9.0×10^{-3}	5.9×10^{-2}	9.5×10^{-6}

The results in Table B.5 indicate that under baseline conditions, the risk was (just) above the average allowable daily probability of infection of $P_{inf} = 2.6 \times 10^{-6}$; under each of the modelled peak conditions, the benchmark was considerably exceeded. The combination of high pathogen loading, poor plant performance and high consumption led to an increase in risk of nearly 4 orders of magnitude. It is therefore relevant to ask: for what proportion of time are baseline conditions representative? How much of the time would the target be expected to be exceeded? What duration of peak conditions could be tolerated to still meet the target of 1×10^{-6} DALY pppy?

B5 Part D: Accounting for variability – stochastic simulation

Rather than calculating discrete point estimates of risk, it is possible to calculate the probability distribution of infection risk. This probability distribution takes into consideration the range of likely values of risk and the probability of each of those values occurring.

B5.1 Data collection

- **Source:** The raw counts (as opposed to only the inferred concentrations; see Annex C, section C2) of *Cryptosporidium* oocysts from the source water samples (n = 10) were obtained.
- **Treatment:** Counts of aerobic spores before and after conventional treatment were obtained from the water treatment facility (n = 8).
- **Exposure:** The original data from the unboiled water consumption survey were obtained.

B5.2 Quantification of model inputs

- **Source:** A negative binomial distribution was fitted to the counts in order to obtain the gamma distribution parameters describing concentration (following the approach presented by Teunis et al. [1997]). The gamma distribution is useful for describing concentration because it must be positive (domain > 0) and flexible in shape, and, when combined with the Poisson distribution to describe counts, it yields a form of the negative binomial distribution (see Annex C, Box C.2). The gamma distribution describing variability in concentration is illustrated in Fig. B.2; the 50th percentile was 0.0026 oocyst·L^{-1}, and the 95th percentile was 1.2 oocysts·L^{-1}.
- **Treatment:** The beta distribution was selected to describe variability in probability of passage (π). The beta distribution is useful for describing probabilities, as it can be defined to have a value only between 0 and 1. The beta distribution plotted on a $\log_{10}$ scale fitted to the aerobic spore counts is illustrated in Fig. B.3; the 50th percentile was a 2.2 $\log_{10}$ reduction, and the 95th percentile was a 1.2 $\log_{10}$ reduction.
- **Exposure:** A lognormal distribution was fitted to the volume of water (mL) consumed per person per day. This distribution is illustrated in Fig. B.4. The median of the distribution was 0.15 L, with the 95th percentile at 0.85 L.

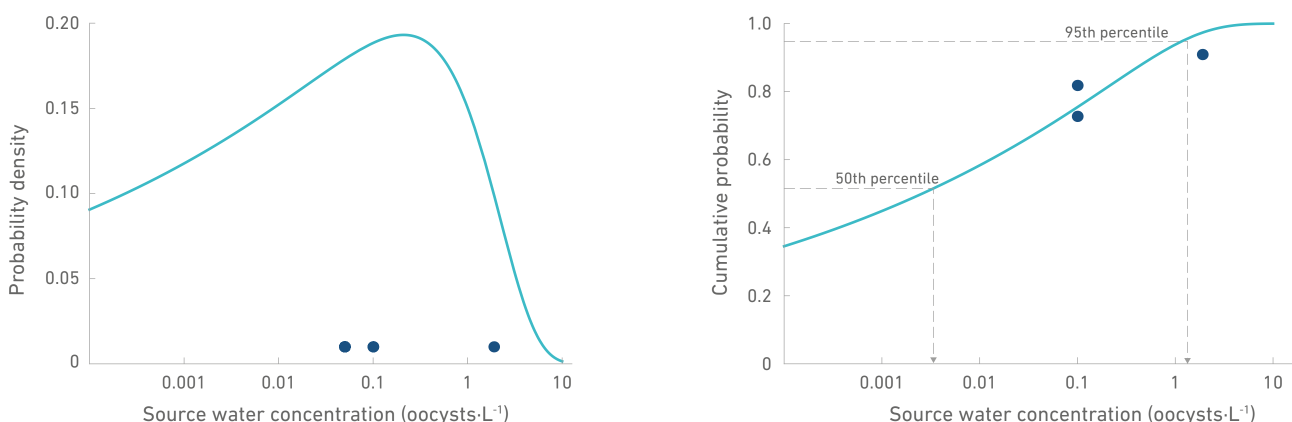


Fig. B.2 Distribution for source water *Cryptosporidium* concentration (oocysts·L^{-1}). Probability density function (left) and cumulative density function (right); positive data points shown.

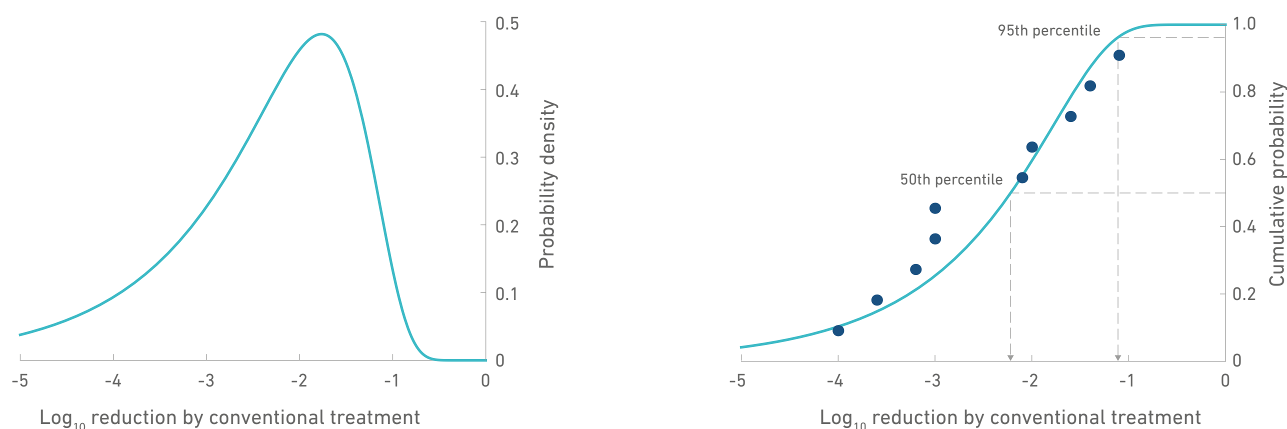


Fig. B.3 Distribution for treatment performance of conventional treatment for aerobic spores. Probability density function (left) and cumulative density function (right). Dots represent paired removal of aerobic spores.

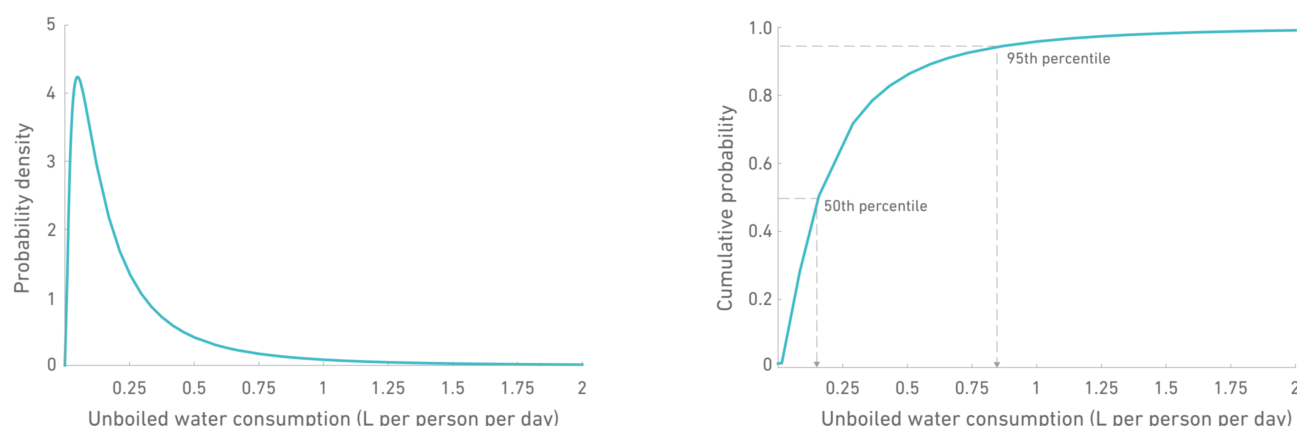


Fig. B.4 Distribution for unboiled water consumption. Probability density function (left) and cumulative density function (right).

B5.3 Risk characterization

These distributions were combined by simulation using Monte Carlo analysis to describe the distribution for the daily probability of infection. The cumulative distribution of risk is illustrated along with the point estimates from sections B2, B3 and B4 in Fig. B.5. The calculated cumulative density function (CDF) indicates that under the assumptions of the model, the health-based target would be exceeded around 55% of the time.

B6 Part E: Accounting for parameter uncertainty – stochastic simulation

In section B5, parametric distributions were fitted to the monitoring data. In most cases, however, the data sets were small, and therefore there was uncertainty associated with the selection of suitable parameter values – which means that there was uncertainty in the shape of the distribution. This uncertainty influences the estimates of risk. The result in Fig. B.5 is the distribution of risk based on combining the most likely shape of each input distribution; however, it would be informative to know how sure we can be about that distribution, given the observed data and the selected models.

B6.1 Data collection

The data collected were the same as for section B5.

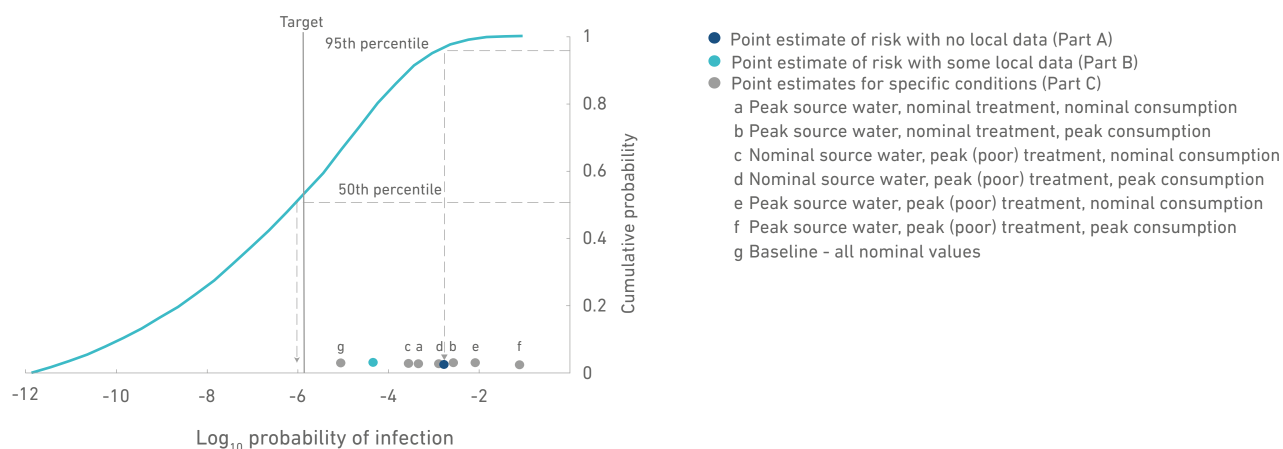


Fig. B.5 Cumulative probability distribution for the probability of infection from Monte Carlo simulation with distributions for source water concentration, treatment performance and unboiled water consumption. Point estimates calculated in sections B2–B4 are also shown.

B6.2 Quantification of model inputs

The uncertainty associated with the selection of parameter values for describing source water concentration, treatment performance and dose–response was explored using Markov chain Monte Carlo sampling of the likelihood function. The 95% credible intervals around the shape of the gamma distribution fitted to the *Cryptosporidium* counts from source water (n = 10), the ($\log_{10}$-transformed) beta distribution fitted to the aerobic spore removal data (n = 10) and the dose–response relationship are illustrated in Fig. B.6.

B6.3 Risk characterization

A second-order Monte Carlo analysis was undertaken; random samples of parameter values were selected, followed by random samples from each model input distribution, and then the risk was calculated. This was repeated thousands of times to obtain a representative sample of the probability of infection. This distribution is illustrated in Fig. B.7. While the best estimate of risk implied that the target would be exceeded approximately 70% of the time, including the parameter uncertainty means that this could be as high as 85% or as low as 35%.

B7 Accounting for other sources of uncertainty

The credible intervals around the CDF for the daily probability of infection in Fig. B.7 take into consideration only the parameter uncertainty surrounding the prediction of the gamma distribution for source water concentration, the beta distribution for treatment performance and the Beta-Poisson dose–response relationship. What about other sources of uncertainty? There are numerous other sources of uncertainty that are not accounted for in those predictions; for example:

- The source water concentration was based on a small monitoring data set collected over a period known to be particularly dry, and rainfall events during that time were small in comparison with the average year. How well can the distribution fitted to those data points be expected to reflect the expected long-term behaviour of the system?
- The distribution system has been ignored; however, the impact of distribution on the consumer risk is an important source of uncertainty.
- Although the water consumption study indicated a relatively low level of water consumption (mean = 0.15 L), some individuals in the community may drink more water, and this would affect the risk.
- The dose–response relationship was developed based on data from a clinical study in which healthy adult volunteers were exposed to laboratory strains of *C. parvum* oocysts. Individuals more susceptible to *Cryptosporidium* may have a higher risk of infection; indeed, *C. hominis* may be more infectious than this relationship describes.

Fig. B.6 a) CDF of gamma-distributed *Cryptosporidium* concentration in source water; b) beta-distributed removal by conventional treatment; c) and dose–response relationship with 95% credible intervals from Markov chain Monte Carlo analysis

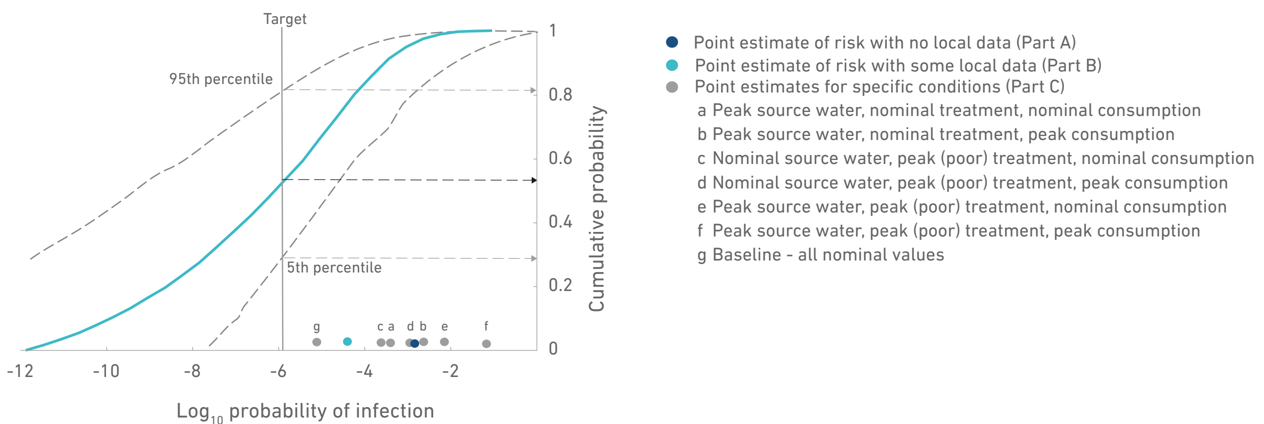


Fig. B.7 CDF of daily probability of infection with 95% credible interval from second-order Monte Carlo simulation

Relying on a sophisticated statistical analysis without holistically thinking about data inputs and assumptions would be an inadequate evaluation of system risks. Quantitatively accounting for these types of uncertainty is a challenge; however, a transparent approach to scenario analysis with point estimations provides a useful tool.

The four issues identified could be addressed using point estimates, by conservatively selecting input values to represent the conditions of concern. For example:

- **Source water concentration:** Data available over a longer period from another climatically relevant site could be analysed to compare the longer-term average and peak concentrations with those in the single "dry" year. Based on this analysis, the input value for source concentration (or the entire distribution) could be conservatively increased.
- **Water consumption data:** To consider events not captured by the local data set, a review of consumption data from the broader international literature indicates that an upper limit of unboiled water consumption would be around 4.5 L per person per day (e.g. Roseberry & Burmaster, 1992). The calculations could be re-run in parallel with this higher value for water consumption, and the sensitivity of the water management outcomes to this input value could be evaluated.
- **Distribution:** Quantifying the impact of distribution systems on infection risk is a challenge; however, scenarios of intrusion events could be simulated to investigate the sensitivity of the model to the distribution network (e.g. van Lieverloo, Blokker & Medema, 2007; Teunis et al., 2010; Besner, Prevost & Regli, 2011).
- **Dose–response:** The outcome could be compared for different dose–response models, including the model fit to *C. hominis* data. In addition, when using single-hit models based on Poisson dispersed microbes (see Annex D), such as the exponential or Beta-Poisson model, the upper limit of infection probability is the probability of exposure. This is represented by the maximum risk curve (exponential model with $r = 1$). The maximum risk curve is a useful upper bound on the uncertainty of the dose–response relationship and can be used for susceptible subgroups or immunocompromised individuals.

In each case, these scenario calculations should be run in parallel with the "best" estimate calculations – they should not replace the best estimates. Using the upper limit of uncertainty at every stage of the model would provide a risk estimate that is unmanageably conservative and not truly representative of the population, and most likely not very helpful for risk management. Alternatively, comparing uncertainty scenario results with the best estimates can provide useful inputs regarding model sensitivity and robustness.

B8 References

- Besner MC, Prevost M, Regli S (2011). Assessing the public health risk of microbial intrusion events in distribution systems: conceptual model, available data, and challenges. Water Res. 45(3):961–79.
- Hijnen WA, Medema G (2007). Elimination of micro-organisms by water treatment processes. KWR Watercycle Research Institute.
- Medema GJ, Teunis PF, Blokker M, Deere D, Davidson D, Charles P et al. (2009). Risk assessment of *Cryptosporidium* in drinking water. Geneva: World Health Organization (http://apps.who.int/iris/bitstream/10665/70117/1/WHO_HSE_WSH_09.04_eng.pdf, accessed 9 April 2016).
- Roseberry AM, Burmaster DE (1992). Lognormal distributions for water intake by children and adults. Risk Anal. 12(1):99–104.
- Teunis PFM, Chappell CL, Okhuysen PC (2002). *Cryptosporidium* dose response studies: variation between isolates. Risk Anal. 22(1):175–83.
- Teunis PFM, Medema GJ, Kruidenier L, Havelaar AH (1997). Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. Water Res. 31(6):1333–46.
- Teunis PF, Xu M, Fleming KK, Yang J, Moe CL, LeChevallier MW (2010). Enteric virus infection risk from intrusion of sewage into a drinking water distribution network. Environ Sci Technol. 44(22):8561–6.
- van Lieverloo JH, Blokker EJ, Medema G (2007). Quantitative microbial risk assessment of distributed drinking water using faecal indicator incidence and concentrations. J Water Health. 5(Suppl 1):131–49.
- WHO (in preparation). Guidelines for drinking-water quality, fourth edition incorporating first addendum. Geneva: World Health Organization.

ANNEX C | MICROBIOLOGICAL DATA AND STATISTICAL INFERENCE

Quantitative characterization of environmental variables based on scientific observations is a critical component of QMRA. For microbial data sets, a background understanding of the methods behind the reported values and an appreciation of the inherent uncertainties in enumeration are useful for developing representative quantitative estimates. There are different possible ways to use these data for statistical inference, each of which may lead to a different quantitative value and, hence, outcome to the QMRA. It is therefore important to ensure that the statistical approach is appropriate to the data and that the implications of any simplifications or data substitutions are understood. In this annex, a brief overview of microbial methods and data analysis considerations is provided. Some examples demonstrating the influence of statistical approach on the quantitative estimates for QMRA are then given.

C1 Microbial enumeration methods

The approach for enumerating pathogens and indicators from environmental samples varies between organisms and between environmental media. Appropriate interpretation of the reported results requires an appreciation of the multiple steps involved in the concentration and purification of samples; the characteristics of the organism that the enumeration method is designed to target; and the quantitative or qualitative nature of the observations.

C1.1 Multiple steps for concentration and purification

Environmental samples need to be prepared for microbial analysis. The method used for preparation will depend on the type of sample (e.g. drinking-water, surface water, sewage, faeces), the expected concentration of organisms present (whether dilution or concentration is required) and which organism is to be targeted. This brief overview draws on the more detailed text by Maier, Pepper & Gerba (2009). Each step in the process is an opportunity for organisms or nucleic acid to be lost; quantitative method controls are required to evaluate each step of the process. Common methodological protocols for different microbial groups are illustrated in Fig. C.1 and described briefly below:

- **Bacteria:** Preparation for bacterial analysis typically involves either direct plating or membrane filtration. In direct plating, the sample (or dilutions of the sample) is added to petri dishes containing growth media and incubated. In membrane filtration (typically used for larger volumes with lower concentrations), the sample is filtered through a membrane (selected depending on target bacteria, but often nitrocellulose filter of pore size 0.45 μm), which is then placed on growth media and incubated.

- **Viruses:** Viruses are often present in low concentrations in surface water samples, and therefore large volumes (up to 1000 L) need to be concentrated. Owing to the small size of virus particles, filtration relies on adsorption of viruses to charged filter media. Viruses are then eluted using fluid (1–2 L) at high pH, which increases the negative charge on both the viruses and the filter, leading to desorption. Further concentration is required to reduce the volume to 20–30 mL before assay, which can then be used for cell culture. If the target virus is to be identified by polymerase chain reaction (PCR), further processing of the sample is required. The nucleic acid is extracted by destruction (lysis) of the cell wall (efficient lysis is a critical step in the process). The sample then needs to be purified from compounds that may interfere with or inhibit the molecular analysis (various methods are used, including centrifuge and filtration); purification methods can lead to the loss of some nucleic acid. For RNA viruses (including norovirus, rotavirus), the RNA must be converted to complementary DNA (cDNA) by reverse transcriptase prior to PCR amplification (reverse transcription PCR, or RT-PCR).
- **Protozoa:** Protozoa are also often present in low concentrations, and therefore large volumes often need to be concentrated. Filtration methods rely on capturing the (oo)cysts by their size. The (oo)cysts are then eluted from the filter using a small volume of elution buffer and shaking for 5 minutes. The protozoa are then pelleted by centrifuge and resuspended in buffer. Immunomagnetic separation is used to separate the (oo)cysts from the debris in the sample. In this process, cysts and oocysts attach to specific antibodies that are associated with magnetic beads and removed from solution. (Oo)cysts are then dissociated from the beads, stained with fluorescent monoclonal antibodies and viewed with an epifluorescence microscope. (Oo)cysts identified by their fluorescence and morphology are referred to as immunofluorescence assay (IFA) positive.

Procedures for sample processing and concentration for pathogen detection

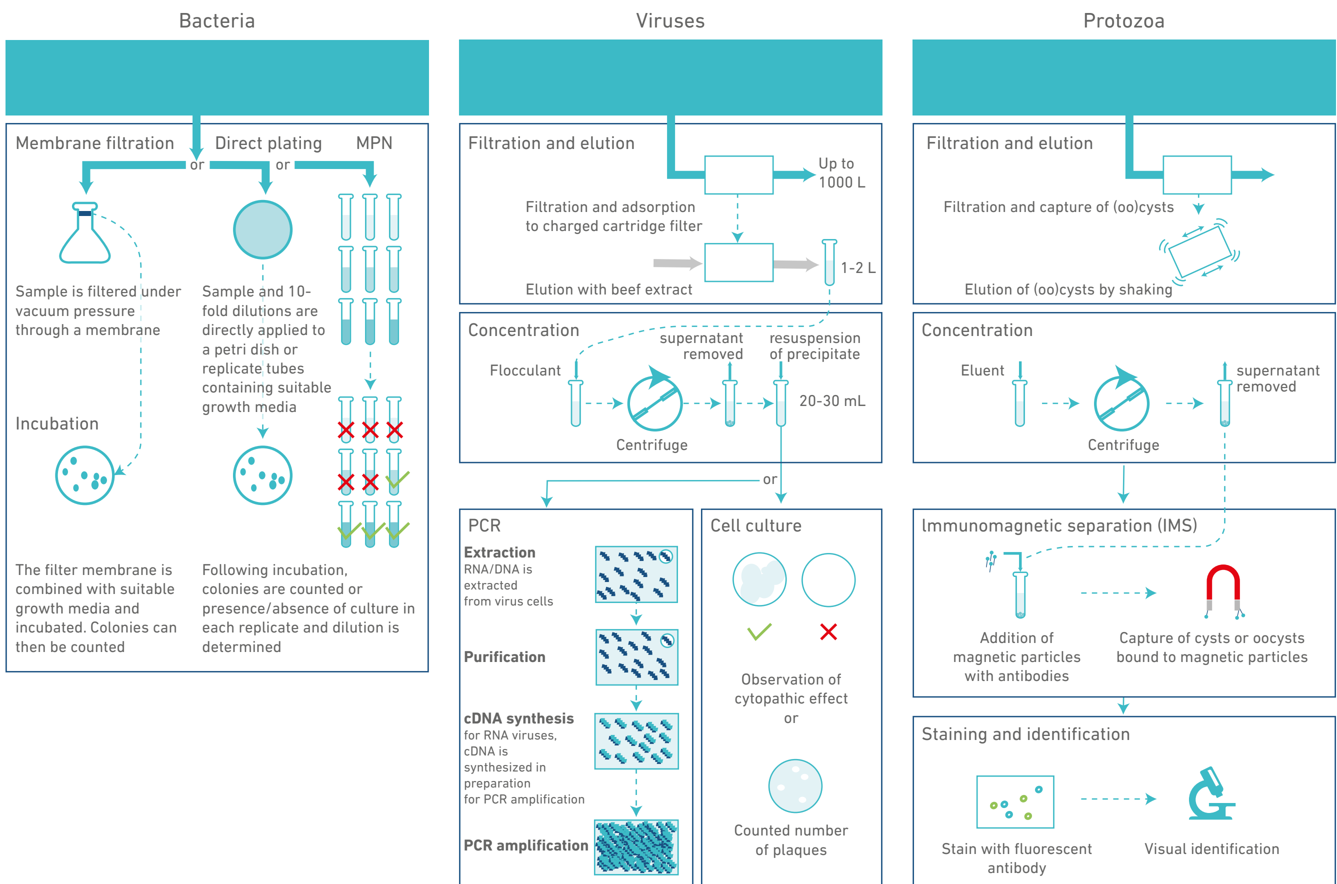


Fig. C.1 Simplified overview of multistep procedures required for sample preparation and enumeration of pathogens

C1.2 Characteristic of the organism that is being targeted

Microbial methods are targeted towards a specific characteristic of the microorganism and vary in their specificity for identifying viable (capable of causing infection), human-infectious organisms. Three dominant methods include:

1) **Visual identification** under the microscope based on characteristic morphological features (often using specific staining techniques). Visual identification methods can verify the integrity of the cell wall and morphological features consistent with a viable organism, but cannot confirm whether an organism is viable or not.
2) **Ability to reproduce** (culture or plaque assay) under a set of specific conditions. Specific organisms can be selected from a large microbial population by using selective media and selective incubation conditions. For pathogenic bacteria, methods usually involve enrichment (to allow damaged cells time to repair and to increase numbers), followed by culture with selective media and/or conditions; a wide range of variations in media and culture protocols exist. For viruses, the two most common approaches to cell culture are the inoculation of a series of wells containing a host cell line with the concentrated sample and then looking for a cytopathic effect where observable changes in the host cells are identified as a result of virus replication. The second approach is plaque assay, where the sample is poured onto a host cell monolayer; following incubation, the areas of dead cells in the host monolayer are counted. The selection of host cell line will influence which enteric viruses are detected, and confirmation (by molecular or immunological methods) of positive cultures is necessary to determine which virus type caused the cytopathic effect or plaque. Culture-based methods identify infectious organisms only; however, as some microorganisms may be viable but non-culturable (i.e. not able to produce a culture in the laboratory, but still infectious to a human host), they can underestimate the number of viable organisms in the sample. Furthermore, environmental organisms (as opposed to those grown in the laboratory) may struggle to grow under laboratory conditions, and the portion that is viable but non-culturable is not well known.
3) **Molecular methods** (usually with amplification by PCR) are used to identify the presence of a particular sequence of genetic material in the sample. Molecular methods are used for pathogens that cannot be cultured (or are difficult to culture) and are sometimes favoured in comparison with culture or visual identification owing to their specificity and sensitivity. Important drawbacks include the following:
   - Standard PCR techniques cannot distinguish between viable and dead organisms.
   - Connecting quantitative results back to the number of microorganisms present is a challenge and depends on the number of target sequences per microorganism.
   - The specificity of the method for targeting the organism of interest depends on the selected probe or primer – the longer the sequence, the more specific the probe or primer is expected to be.

C1.3 Quantitative or qualitative observations

Although results are often reported as a concentration, because microorganisms are discrete units, the concentration cannot be measured directly and is an inference based on one of three types of observations:

1) **Quantitative** (counted number of organisms, colonies or plaques), where the precision for predicting the original concentration depends on the number of organisms counted and the volume (including dilution) of the sample (e.g. a count of 2 in 100 mL is less precise than a count of 200 in 10 L).
2) **Qualitative** (presence or absence of an observed response in a sample volume). Quantitative concentrations are frequently reported from a series of qualitative results as the MPN. The precision associated with this MPN in quantifying the original concentration depends on the number of observations, which may be small (e.g. three 10-fold dilutions) or large (Colilert® Quanti-Tray®, with 97 observations per sample). The predicted concentration is based on the assumption that organisms are randomly (Poisson) distributed in the sample and is typically reported (in standard MPN tables) with a 95% confidence interval.
3) **Semiquantitative**. Quantitative PCR (qPCR) uses the rate of amplification of the target genetic material in the sample to predict the original concentration of genomes or genome equivalents in the environmental sample. For QMRA, this approach is currently considered semiquantitative, as the translation of the qPCR result back to an absolute concentration of infectious pathogenic units in the original sample is uncertain. As methods for quantifying recovery (loss of organisms and nucleic acid), extraction (efficiency of extracting genetic material from microbial cells), inhibition (potential suppression of the PCR signal) and the relationship between genome units to infectious particles are improved, so too will the quantitative value of qPCR data for QMRA.

C2 Statistical inference

Whereas the observations may be qualitative, semiquantitative or quantitative, monitoring data are typically reported from the laboratory as microbial concentrations (e.g. MPN·100 mL^{-1}, CFU·100 mL^{-1}, PFU·100 mL^{-1}). In order to reach these concentrations, the laboratory applies a set of standard protocols (which vary between organisms, methods and laboratories) for inferring the concentration from its observations. Quantitative inference for model

inputs can be undertaken based on the reported concentrations or on the original observations; however, the adopted approach has implications for the uncertainty that is explicitly represented and accounted for in the final risk assessment. The different levels of statistical inference (Box C.1) are illustrated in Fig. C.2.

Box C.1 Statistical inference

Statistical inference refers to the process of drawing conclusions from the data. In contrast to description statistics, where the purpose is to quantitatively describe a set of data, statistical inference is a systematic approach used to draw conclusions from data sets arising from systems affected by random variation.

Statistical inference will most often use:
- a model used to describe the random process; or
- observational data.

The conclusion of the statistical inference could include:
- an estimate of a parameter value of the model (e.g. the mean concentration; or two parameters describing the gamma distribution); or
- a confidence interval (frequentist) or credible interval (Bayesian) around the parameter values.

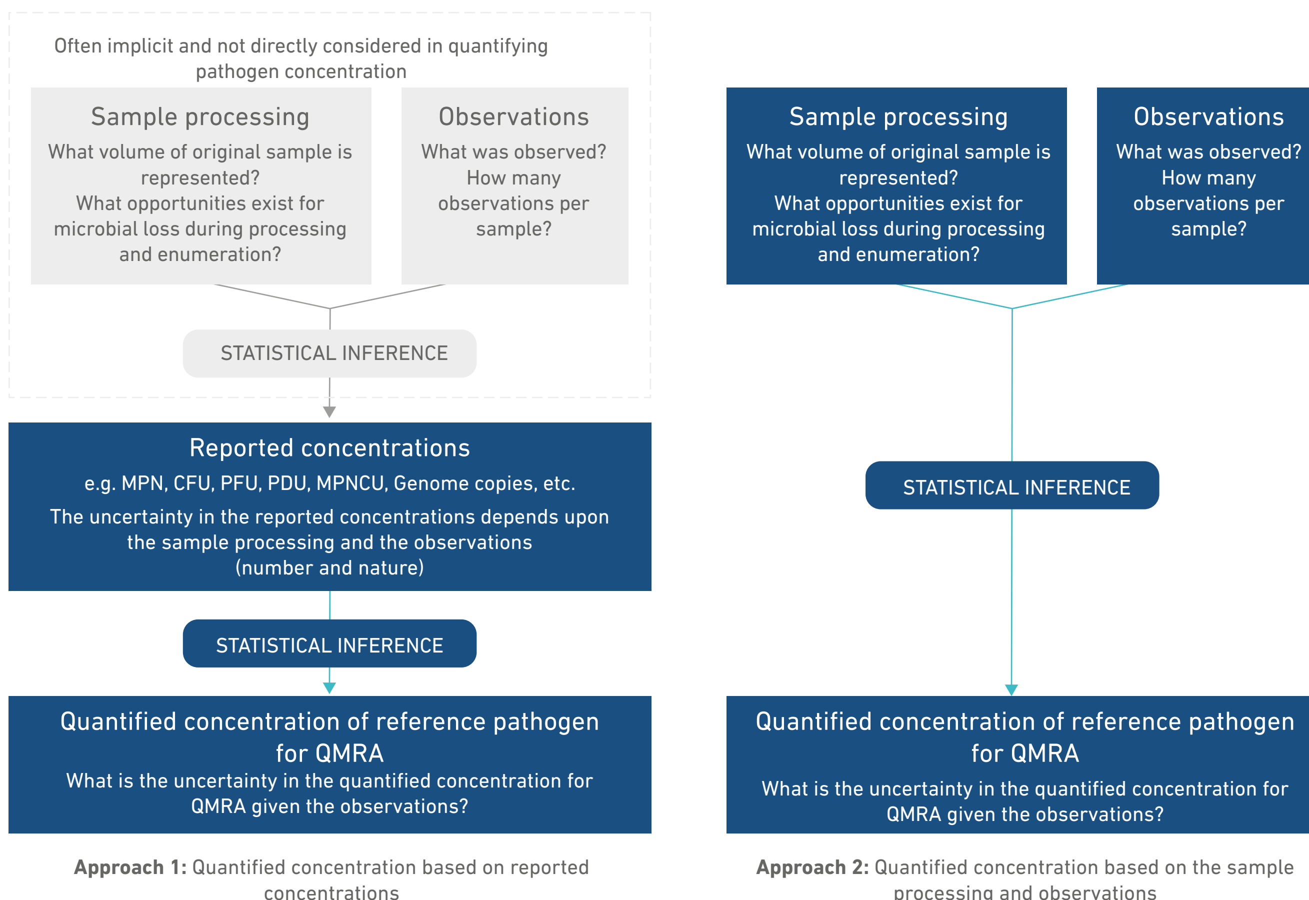


Fig. C.2 Two approaches to the statistical inference of pathogen concentration based on either reported concentrations or original observations. MPNCU: most probable number of cytopathic units; PDU: PCR detectable units.

C2.1 Inference based on reported concentrations (Approach 1 in Fig. C.2)

Reported concentrations are themselves a statistical inference based on the laboratory observations. The certainty (precision) associated with this concentration estimate varies between organisms and depends on the steps and

unquantifiable loss involved in sample preparation; the characteristics of the microorganism that is targeted by the method (how specific the method is to the organism of interest); and the nature (count, presence/absence) and number of the observations associated with each sample. This is illustrated conceptually in Fig. C.3.

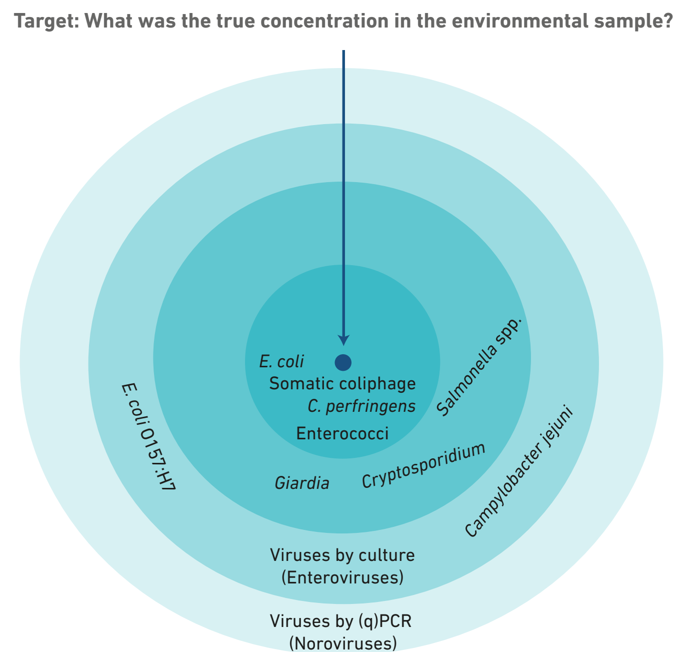


Fig. C.3 Relative precision associated with different methods of enumeration for predicting pathogen and indicator concentration in environmental media

Concentration predictions based on indicator organisms are the most precise. They are selected as indicators because of the relative simplicity (in comparison with pathogens) of their methods, and hence there are typically a larger number of observations associated with each concentration prediction (e.g. Colilert® Quanti-Tray®). *Giardia* and *Cryptosporidium* are included in the next tier, as there are well-characterized protocols for measuring the recovery efficiency of the methodological process; the results are quantitative (i.e. counts); and there are standard methods, leading to greater reproducibility between laboratories.

Salmonella is also included in this tier, owing to limited sample processing (and potential for loss), a well-defined MPN matrix approach for quantification and the existence of well-established standard methods. Predictions based on the other pathogenic bacteria (*Campylobacter* and *E. coli* O157:H7) are less precise, as the organisms are more difficult to culture, it is more difficult to identify the target organism, and MPN results are often not standard matrices and consist of fewer observations than the 3 × 3 or 5 × 5 tables. Virus enumeration by culture is included in the same tier, with concentrations based on the MPN procedure or plaque counts; recovery is rarely accounted for in viral assays. Results from molecular methods are the least precise, as a result of unquantifiable loss of genetic material, unknown performance of extraction, unknown magnitude of dampening of PCR signal, unknown translation of genetic copies to virus particles and the lack of knowledge regarding viability.

For statistical analysis, the uncertainty associated with the reported concentrations increases with distance from the centre of the circle, and hence the implications (in terms of unaccounted for uncertainty) for the final concentration predictions also increase with distance from the centre. For example, reported *E. coli* concentrations using Colilert® Quanti-Tray® are very precise, and direct analysis of these concentrations is reasonable; however, simply relying on reported virus concentrations (by culture or PCR) is likely to significantly underestimate the uncertainty associated with the final concentration estimate.

C2.2 Inference based on original observations (Approach 2 in Fig. C.2)

Statistical inference of the pathogen concentration for QMRA based directly on the laboratory observations (counts, presence/absence in a given sample volume) is theoretically preferred, because the assumptions and uncertainties can be more explicitly incorporated into the analysis. The mathematical complexity associated with the analysis

is, however, increased, and appropriate statistical modelling is important. Particularly for drinking-water studies, quantification of concentration for QMRA is increasingly being based on the actual observations. Many examples exist for protozoa, for which counts and volumes have been analysed using discrete distributions, including the Poisson distribution (Makri, Modarres & Parkin, 2004; Åström et al., 2007), the negative binomial distribution (Teunis et al., 1997; Masago et al., 2002, 2004; Pouillot et al., 2004) and the Poisson lognormal distribution (Masago et al., 2004; Signor & Ashbolt, 2006). For presence/absence data, concentration estimates have been based directly on the observations, in many cases including uncertainty bounds (Westrell et al., 2006; Åström et al., 2007; Petterson et al., 2009).

There are, however, two potential complications associated with this approach:

1) **Obtaining the original observations:** Owing to standard reporting protocols at laboratories, the original observations are usually not included in electronic databases and therefore are often not easily accessible. Obtaining the original results (e.g. total sample volume, volume of sample processed, presence/absence per tube, count per plate, etc.) may require reference to handwritten laboratory reports.
2) **Mathematical complexity:** Constructing statistical models that account for the specific nature of the observations requires a level of mathematical skill that is often beyond the risk assessor. Contracting specialist skills may be of value and should be considered, depending on the scope and purpose of the assessment.

C3 Quantifying pathogen concentration

C3.1 Data considerations

The following factors need to be considered when interpreting pathogen data:

- **What type of data?** The data may be qualitative (presence or absence in a given volume); discrete (only contains integer [0, 1, 2, 3…]; categorical (can hold one of a range of distinct values, as in the case of MPNs); or continuous (can hold any value within a set range, as in the case of reported concentrations). It is important to select distributions or approaches that are appropriate to the type of data (see Box C.2).
- **What to do about zeros?** Microbial count data often include (sometimes many) zeros. When using a discrete distribution to describe concentration (e.g. Poisson, negative binomial; see Box C.2), the zeros are no problem, as a zero is within the domain of the distribution. However, there are many examples in the literature where continuous distributions have been used to describe discrete microbial counts (Masago et al., 2002, 2004; Smeets et al., 2007; Jaidi et al., 2009). In order to do this, counts must be converted to concentrations (concentration = count/volume), and zeros need to be replaced with an alternative assumption. The most common approaches include replacing zeros with the detection limit, with half the detection limit or with a continuous distribution between zero and the detection limit. Two studies compared the impact of different zero replacement assumptions on their risk estimates (Smeets et al., 2007; Jaidi et al., 2009) and concluded that assumption choice strongly influenced the risk estimates when overall pathogen concentrations were low (Smeets et al., 2007; Jaidi et al., 2009) and when data sets were small (Jaidi et al., 2009). The replacement can influence the final concentration estimation, as illustrated in section C3.4.1; therefore, whatever method is applied, the impact on the concentration and the subsequent risk assessment results should be tested. This can be done within sensitivity analysis (see Chapter 8) by running both options and comparing the impact on the risk management outcome of the risk analysis.
- **What to do about censored data?** Reported microbial concentrations typically contain left-censored (less than) and right-censored (greater than or TNTC) data. As for the specific case of zero counts, it is common to replace left-censored data with the detection limit, with half the detection limit or with a continuous distribution between zero and the detection limit. Similarly, right-censored values are often replaced with the upper limit of detection or twice the upper limit of detection. The suitability of replaced values depends on the individual data set and the analytical method that has been applied. It is possible to construct likelihood functions for fitting distribution parameters using the actual observations of less than or greater than results. It is advisable to investigate the sensitivity of the risk assessment results to assumptions applied within risk characterization (Chapter 8) by running several options and comparing the results.

Box C.2 Stochastic simulation

Probability distributions

A probability distribution can be used to describe how a model input will vary: the range of expected values, and the probability that the input is equal to or less than those values. The probability distribution goes beyond simple descriptors (mean, median or upper 95th percentile) and characterizes the entire distribution.

The cumulative distribution function (CDF) or probability distribution function (PDF) is the mathematical equation that describes the probability that a variable X is less than or equal to x. That is,

$$F(x) = P(X \leq x) \text{ for all } x \qquad \text{Eq. C.1}$$

A PDF $f(x)$ is the derivative of the CDF and describes how the probability mass varies over the range of the variable. The PDF is defined as:

$$f(x) = \frac{d}{dx} F(x) \qquad \text{Eq. C.2}$$

Parametric functions are often used to define probability distributions for risk analysis. The selected distribution will depend on the domain of the model input (what values are theoretically possible) and how it is expected to vary. Some commonly used distributions in QMRA include:

- **Poisson:** A discrete distribution that is often used to describe microbial counts [Domain: x = 0, 1, 2, 3, ...].
- **Negative binomial:** A discrete distribution that can be used to describe overdispersed (clumped) microbial counts [Domain: x = 0, 1, 2, 3, ...].
- **Gamma:** Often used for describing variability in pathogen concentration due to a convenient relationship. When describing microbial counts with a Poisson distribution, if the Poisson parameter is assumed to vary according to a gamma distribution, the resulting count distribution is negative binomial [Domain: $x \geq 0$].
- **Lognormal:** Often used to describe microbial concentrations or exposure volumes [Domain: $x \geq 0$].
- **Beta:** Can take a wide range of shapes and is often applied to describe probabilities (e.g. probability of passage for reduction barriers; probability of recovery by an analytical method) [Domain $0 \leq x \leq 1$].

Monte Carlo simulation

When each model input is described by a probability distribution, quantifying the exact mathematical distribution for the risk output (e.g. probability of infection) can be complicated or (depending on the input distributions) impossible. Monte Carlo simulation is a random sampling technique that allows the outputs of the risk model to be quantified by sampling. Random samples are selected from each input variable distribution, and the risk output is quantified; this is then repeated thousands of times to obtain a random sample of the risk output. The frequency distribution of output samples is assumed to represent the probability distribution of risk. Typically, in QMRA models, all input variables are assumed to be independent. It is important to recognize that this may not be the case, and Monte Carlo simulation can overinflate the expected variability and associated uncertainty in the risk outcomes.

For more details on distributions, see Vose (2008).

- **How to analyse MPNs?** Reported concentrations based on MPN inference are predicted concentrations from presence/absence data and frequently include left- and right-censored outcomes. MPNs are often assumed to represent continuous data; however, when based on a standard MPN table (e.g. 3 × 3 or 5 × 5), the data are categorical (only a discrete number of outcomes are possible, depending on the presence/absence observations). The importance of this simplification depends on the context of the risk assessment and the accuracy required for the statistical analysis. This is illustrated in section C3.4.2 for *Campylobacter* enumeration data from surface waters.

C3.2 Accounting for method recovery

When method recovery is not accounted for in the concentration predictions, the concentration will be underestimated. Although recognized as important, quantitative incorporation of method recovery depends on the

availability of observations to describe the recovery. When recovery data are available for analysis, it is necessary to consider the following:

- It is important to know whether the recovery experiments were undertaken using the same sample. Recovery performance is influenced by the sample matrix, and therefore it is important to ensure that the medium used in the recovery experiments is comparable with the true environmental samples. Sample-specific internal recovery spikes (e.g. ColorSeed™ for *Cryptosporidium* and *Giardia*) are ideal in this regard, as the measure of recovery relates to each specific sample.
- For culture methods, recovery experiments can be influenced by the type, source and culture conditions of the surrogate microbes. For example, laboratory strains of bacteria cultured on artificial media and then used in recovery efficiency studies of spiked samples may differ in recovery efficiency from bacteria present in environmental samples that were excreted in faecal matter and subjected to environmental stressors that induced injury. Such differences in physiological states of microbes in environmental samples compared with laboratory-grown positive control microbes used in recovery efficiency studies can produce dramatic differences in microbial recovery efficiencies. Such differences in recovery efficiency have been well documented in the food microbiology literature and are often not considered or adjusted for in the methods used for the analysis of bacteria in water and wastewater.
- Which part of the overall analytical sequence is the recovery data describing? Recovery data may be available for some parts, but not all, of the overall analytical sequence. For example, van Heerden and coworkers (2005) assumed that adenovirus recovery by PCR was 40%, based on published recovery rates of the adsorption–elution method adopted using glass wool.

There are a considerable number of data describing method recovery for *Cryptosporidium* and *Giardia* where (oo)cysts are spiked into the original sample (either a parallel sample or the same sample using labelled oocysts such as ColorSeed™) and then analysed alongside the native (oo)cysts. Most approaches suffer from the lack of a "gold" standard, in that the formulation of the spike typically presents microbes in a "cleaned-up" or freshly cultured state, as well as in a relatively high density, both of which may have an impact on how a method recovers them in comparison with native target microbes in an environmental matrix. Nonetheless, several approaches have been used for incorporating these data into concentration estimates for QMRA:

- **For independent[1] samples:**
    a) The binomial probability of recovery (each organism has a certain probability of being counted) has been assumed to follow a beta distribution and to be included as an independent random variable in the risk model (Teunis et al., 1997; Makri, Modarres & Parkin, 2004; Pouillot et al., 2004; Signor & Ashbolt, 2006). This is the most common approach for incorporating recovery into concentration estimates for QMRA. Adopting a modification, Jaidi and coworkers (2009) also introduced a coefficient of correlation between counts and recovery into their analysis.
    b) Avoiding the need to assume a beta distribution, the recovery data can be incorporated into the concentration estimate by random sampling of the recovery estimates to "correct" each count (Medema et al., 2003).
- **For sample-specific (paired) recovery estimates:** Each count can be directly corrected (Medema et al., 2003; Petterson, Signor & Ashbolt, 2007), reducing the between-sample uncertainty.

There are limited examples of including recovery estimates for viral and bacterial concentration estimates; however, as mentioned previously, van Heerden and coworkers (2005) assumed that adenovirus recovery by PCR was 40%, based on published recovery rates of the adsorption–elution method adopted using glass wool; and Petterson and coworkers (2009) incorporated recovery into predicting *E. coli* O157:H7 concentration in surface water from presence/absence data.

Unfortunately, in many cases, method recovery data do not exist, and the limitation of this data gap for quantifying concentration for QMRA cannot be overemphasized. This limitation is particularly relevant for molecular methods, where many steps are involved in the processing and analysis of samples, and the availability and application of quantitative controls are often limited. If the loss is unknown, then the prediction of the original concentration is impossible, except to say that "at least" the estimated concentration based on the observations was originally present. When no data are available, it is a good idea to conservatively estimate recovery to ensure that the risk is not underestimated.

[1] The recovery results are taken from a set of samples separate from those for which the pathogen was enumerated.

C3.3 Accounting for viability/infectivity and human infectivity

Confusion exists surrounding the terms viability and infectivity. In the context of this report, the following definitions are held:

- **Viability:** obtaining some characteristics (as defined by the viability test) consistent with being capable of causing infection;
- **Infectivity:** capable of causing infection in a host; and
- **Human infectivity:** capable of causing infection in a human host.

The first concern is whether the organism is capable of causing infection, here referred to as viability/infectivity (where viability is an approximation for infectivity). The second concern is whether the organism is human infectious, referring to host specificity and whether the organism is capable of causing infections in humans. Considerations for the quantification of these factors as related to the methods of detection are outlined below:

- **Culture:** For culture-based enumeration methods, the identified organisms are clearly still capable of causing infection (in fact, culture methods are more likely to underestimate the number of infectious organisms, as some organisms may still be infectious to humans yet unable to multiply in artificial growth media or cell culture). Quantitative inclusion of viable but not culturable microbes is possible only when observations are available to describe the culturability.
- **Visual identification:** There is no definitive measure of the viability or infectivity of *Cryptosporidium* or *Giardia* (oo)cysts by visual observation. Measures based on the identification of morphological characteristics consistent with viability have been suggested, including visual identification of sporozoites within the oocysts (diaminophenylindole [DAPI] staining) and verification of cell membrane integrity with vital staining (e.g. propidium iodide; Dowd & Pillai, 1997). These methods are more useful in indicating the presence of dead (oo)cysts in the population counted rather than being a reflection of the infectivity. However, viability based on these visual observations, even with viability staining methods, is not a reliable predictor of human, animal or cell culture infectivity. Data such as these can be included in the concentration estimates in one of two ways. Firstly, only the observed potentially viable organisms can be counted, and hence any apparently dead organisms are discarded; however, not all "viable" cells may be infectious. Secondly, viability can be considered a binomial (two-outcome) process in which each identified (oo)cyst has a certain probability of being potentially infectious (P); and the probability of not being infectious is ($1 - P$). The data can be used to estimate the probability of potential infectivity; in a similar manner to the recovery, the probability of being potentially infectious can be modelled as an independent (possibly beta-distributed; see Box C.2) random variable in the QMRA.

  The above approach was applied by Teunis and coworkers (1997) using data reported by LeChevallier, Norton & Lee (1991). For samples collected across the USA and Canada, approximately 32% of the 242 *Cryptosporidium* observed in raw water samples contained sporozoites within the oocyst (as a measure of viability). The raw data from this study were reanalysed by Teunis et al. (1997) and applied in their QMRA of surface waters in the Netherlands, by fitting a beta distribution to the potentially viable proportion. This same distribution was subsequently applied in studies in Japan (Masago et al., 2002) and the USA (Makri, Modarres & Parkin, 2004). The limitation here is whether the data set collected across the USA and Canada is representative of different settings; viability relates to oocyst persistence, which is influenced by numerous environmental factors, the combination of which is expected to be site specific. Therefore, site-specific consideration of viability is necessary, and when there are concerns that the data are not representative, the importance of viability assumptions to the overall risk estimates should be tested during sensitivity analysis within risk characterization (see Chapter 8); if important, local data should be obtained, or more conservative assumptions adopted.
- **Molecular methods:** Molecular methods such as qPCR target the genetic material of the pathogen and therefore detect both viable and non-viable organisms. Several studies have relied on PCR data for quantifying source material for QMRA, including noroviruses in tap water (Masago et al., 2006); noroviruses and rotaviruses in household water (Oesterholt et al., 2007); noroviruses and enteroviruses in surface water (Åström et al., 2007); and *Cryptosporidium* and *Giardia* in canal water in Thailand (Diallo et al., 2008). In each of these studies, all identified organisms were assumed to be infectious for the purposes of QMRA. There is clear evidence from disinfection studies that all viruses amplifiable by RT-PCR and PCR are not infectious, especially if they have been exposed to disinfectants and other stressors (Sobsey et al., 1998; Ko, Cromeans & Sobsey, 2003). Therefore, molecular methods still typically overestimate infectious organism concentration.

Owing to the limitations in available data, infectivity can rarely be quantitatively accounted for in QMRA. However, whether the identified organisms are still in an infectious state is a critical consideration for assessing the health

implications of a measured concentration. In the absence of data, consideration of the sources and transport of pathogens in environmental samples is important for the assessment of infectivity. For example, it may be reasonable to assume that genetic material sourced from undisinfected wastewater may still represent infectious viruses (Åström et al., 2007). However, norovirus RNA in disinfected drinking-water is much less likely to be infectious (Masago et al., 2006).

Some authors have argued that recovery and infectivity are of similar magnitude and therefore can be assumed to cancel each another out (Regli et al., 1991; Smeets et al., 2007). Despite the pragmatic advantages, there is no real scientific justification for this approach.

C3.4 Example: modelling the distribution of concentration from pathogen data

Samples were collected from a surface water source impacted by agriculture and analysed for *Cryptosporidium* and *Campylobacter*. The reported concentrations are included in Table C.1. The objective of the analysis is to characterize the probability distribution of concentration for each pathogen in the surface water source for QMRA. The following example demonstrates the practical interpretation of the pathogen data sets and illustrates the implications of different statistical analysis approaches on quantitative concentration estimates.

Table C.1 Reported concentrations of *Cryptosporidium* and *Campylobacter* for a surface water source

Sample number	*Cryptosporidium* (oocysts·L^{-1})	*Campylobacter* (MPN·L^{-1})
1	<0.3	1 100
2	0.23	90
3	<0.43	20
4	0.32	>2 400
5	<0.062	500
6	<0.025	90
7	<0.056	20
8	<0.05	500
9	0.19	20
10	<0.28	21
11	<0.26	500
12	<0.013	20
13	0.37	90
14	0.043	500
15	<0.13	200
16	0.85	23
17	1.21	>2 400

C3.4.1 *Cryptosporidium*: analysis of microbial count data

The objective of the analysis was to fit a PDF for oocyst concentration in surface water to the data. Some factors associated with the reported *Cryptosporidium* concentrations were unclear:

- The method for detection of *Cryptosporidium* is to count the number of oocysts in the volume of sample filtered. The reported concentrations of *Cryptosporidium* in Table C.1 are therefore an inference based on the number of organisms counted in each sample volume (see section C1.2).
- It is not clear if the counts were total IFA positive oocysts or confirmed by DAPI (see section C1.3).
- Was the method recovery included in the calculation of the reported concentrations (see section C1.2)?

To assist with analysing the data, it was necessary to approach the laboratory to obtain the counts and volumes behind the reported data, information on the methodology and the results from recovery seeding. The observations behind the reported concentrations were obtained and are included in Table C.2.

Table C.2 *Cryptosporidium* sampling data from a surface water source

Sample number	Count (IFA + oocysts)	Volume (L)	Recovery[a]	Daily concentration estimates[b] (oocysts·L^{-1})
1	0	10	33	<0.3
2	1	10	44	0.23
3	0	5	47	<0.43
4	1	12.5	25	0.32
5	0	43.4	37	<0.062
6	0	83.3	48	<0.025
7	0	33.3	54	<0.056
8	0	62.5	32	<0.05
9	1	11.1	48	0.19
10	0	12.5	29	<0.28
11	0	11	34	<0.26
12	0	125	62	<0.013
13	2	10	54	0.37
14	2	113.6	41	0.043
15	0	15.4	49	<0.13
16	3	12.2	29	0.85
17	14	32.1	36	1.21

[a] Number of labelled oocysts recovered from 100 spiked.

[b] The daily concentration estimate was given by: (Count/Volume)/(Recovery/100). If the count was zero, then the detection limit was determined by setting the count equal to 1.

A gamma distribution was selected to describe the oocyst concentration because of the relationship between the negative binomial count distribution and the gamma distribution for concentration (see Box C.1). Three different analysis approaches were applied to the same data set in order to show the difference in the estimated distribution for describing concentration when the reported concentrations were used versus the observations of counts and volumes; and, in particular, the impact of zero replacement assumptions (see section C1.1):

1) A negative binomial (Poisson/gamma mixture) distribution was fitted to the reported counts and the volumes. *A discrete distribution fitted to observations.*
2) A gamma distribution was fitted to the daily concentration estimates (n/V), with zeros substituted with the detection limit. *A continuous distribution fitted to the reported concentrations with zero replacement – option 1.*
3) A gamma distribution was fitted to the daily concentration estimates (n/V), with zeros substituted with half the detection limit. *A continuous distribution fitted to the reported concentrations with zero replacement – option 2.*

In each case, a statistical model was constructed both without accounting for recovery and then with accounting for recovery. The method of maximum likelihood (e.g. Walpole et al., 2011) was used to fit each model to the observed data in order to estimate the parameters of a gamma distribution. The maximum likelihood estimators (MLE) of the gamma distribution parameters, along with the expected value of the concentration and the upper 95th percentile of the MLE distribution, are given in Table C.3. The CDFs from each approach (accounting for recovery) are illustrated with the daily concentration estimates in Fig. C.4.

Table C.3 Results of maximum likelihood estimation of gamma distribution parameters and predicted *Cryptosporidium* concentrations, given different statistical methods

Analysis approach	Gamma parameters		Concentration predictions (oocysts·L^{-1})	
	$\hat{\rho}$	$\hat{\lambda}$	$\hat{E}^a$	Upper 95th percentile
No accounting for recovery				
Discrete 1) Negative binomial distribution	0.25	0.27	0.07	0.47
Continuous 2) Gamma distribution, zero substitution equal to detection limit	1.07	0.10	0.11	0.38
3) Gamma distribution, zero substitution equal to half detection limit	0.79	0.11	0.09	0.36
Accounting for recovery				
Discrete 1) Negative binomial distribution	0.23	0.81	0.18	1.32
Continuous 2) Gamma distribution, zero substitution equal to detection limit	0.95	0.30	0.28	1.07
3) Gamma distribution, zero substitution equal to half detection limit	0.72	0.33	0.24	1.00

[a] Expected value of concentration (mean) is given by $\hat{\lambda}\cdot\hat{\rho}$.

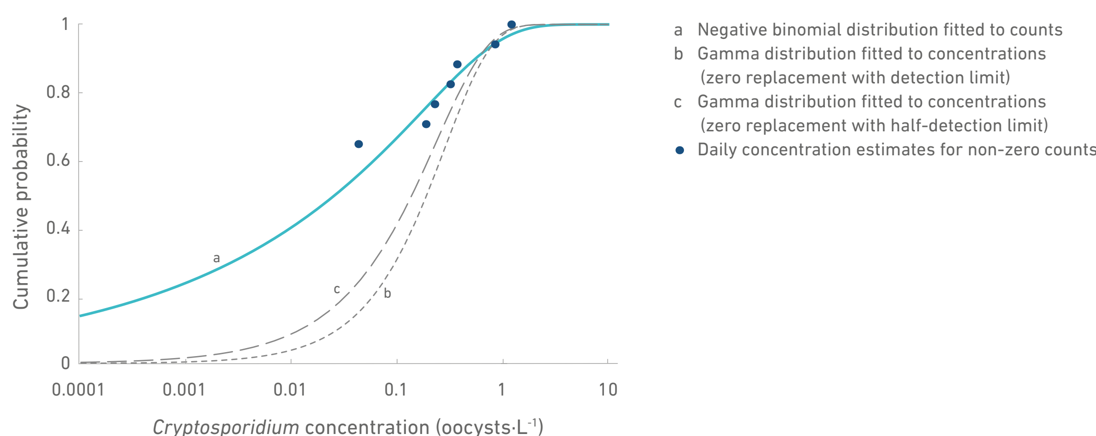


Fig. C.4 Maximum likelihood gamma distributions describing *Cryptosporidium* concentration given different statistical approaches (recovery accounted for within the model)

The following comments relate to the specific pathogen data considerations raised in this section:

- **Reported concentrations versus observations:** Zero substitution led to less variability in the estimated density function, with higher minimum and lower maximum concentrations. The gamma distribution fitted to the actual counts had a lower mean, but higher predicted upper 95th percentile concentration. The importance of the quantitative difference in concentration depends on the application of the QMRA and the level of detail required in the analysis.
- **Accounting for method recovery:** Accounting for method recovery had a clear impact, leading to a 2- to 3-fold increase in the predicted concentrations. When recovery was ignored in the interpretation of pathogen data for QMRA, the concentrations were underestimated.
- **Accounting for viability/infectivity:** The concentrations predicted are for total IFA positive oocysts, and no information is provided on the viability or infectivity of the counted oocysts. Undertaking the QMRA based on the concentration of total oocysts may provide a first, conservative estimate of the risk. If the conservatively estimated risk is of concern, the likelihood of risk of infection should be further considered.

- **Accounting for human infectivity:** The distribution quantifies the total concentration of oocysts and does not distinguish between species, some of which are more infectious to humans than others.

C3.4.2 *Campylobacter*: analysis of presence/absence data

As for *Cryptosporidium*, the objective was to describe the PDF for *Campylobacter* concentration in surface water. Some factors were unclear:

- The reported concentrations of *Campylobacter* were an inference based on the presence/absence results (MPN·L^{-1}) by enrichment and cell culture with selective media and conditions (42 °C). What were the presence/absence observations behind the reported concentrations?
- Was the method recovery accounted for in the reported concentrations?

Further details on the observations associated with the reported concentrations of *Campylobacter* are given in Table C.4. The concentration estimates were the MPN based on a 3 × 3 matrix (three replicates at three 10-fold dilutions). The data were used to fit the parameters of a gamma distribution to describe variability in concentration using two different approaches:

1) The gamma distribution was fitted to the reported concentrations with the right-censored data substituted with the maximum value (2400 MPN·L^{-1}).
2) A hierarchical statistical model was fitted to the presence/absence results for *Campylobacter*. The model was constructed to allow for the daily concentration estimates to be the Poisson parameter and the between-day variability to be described by a gamma distribution.

The results are given in Table C.4 and illustrated in Fig. C.5.

Table C.4 *Campylobacter* sampling data from a surface water source

Sample number	Reported result (MPN·L^{-1})	Observations (number of positives from three replicates in each volume)			Daily concentration estimate (Poisson parameter) (MPN·L^{-1})	Uncertainty quantile (95% credible interval)
		100 mL	10 mL	1 mL		
1	1 100	3	3	2	1 099	[91, 1 138]
2	90	3	2	0	93.3	[14, 180]
3	20	2	1	1	20.5	[4, 48]
4	>2 400	3	3	3	–	[108, 1 640]
5	500	3	3	1	462	[69, 663]
6	90	3	2	0	93.3	[14, 244]
7	20	2	1	1	20.5	[5, 49]
8	500	3	3	1	462	[49, 701]
9	20	2	1	1	20.5	[5, 47]
10	21	2	2	0	21.1	[4, 52]
11	500	3	3	1	462	[55, 672]
12	20	2	1	1	20.5	[5, 51]
13	90	3	2	0	93.3	[13, 212]
14	500	3	3	1	462	[58, 693]
15	200	3	3	0	239.8	[28, 523]
16	23	3	0	0	23.11	[5, 65]
17	>2 400	3	3	3	–	[77, 1 906]

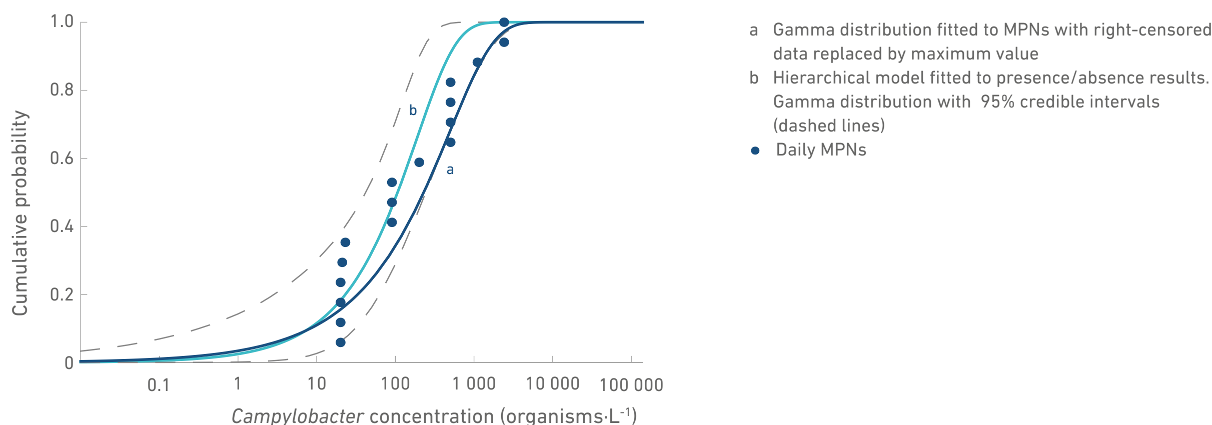


Fig. C.5 Maximum likelihood gamma distributions describing *Campylobacter* concentration given different statistical approaches

The following comments relate to the specific pathogen data considerations raised in this section:

- **Reported concentrations versus observations:** The reported concentrations and the fitted concentrations from the hierarchical model are given in Table C.4 (column 2 and column 6, respectively). The uncertainty in the estimated daily concentration is given by a credible interval in Table C.4 (column 7). The MLE gamma distribution CDF under each approach is illustrated in Fig. C.5. For this data set, using the reported MPNs resulted in a higher, more conservative distribution.
- **Accounting for method recovery:** There was no quantitative information available on the recovery of the method, and therefore recovery has not been accounted for in this distribution.
- **Implications for viability:** The method of identification was by enrichment and culture, and therefore only viable organisms were identified; however, it may be that the predictions are an underestimate of the total *Campylobacter* load, owing to the presence of viable but non-culturable organisms.
- **Implications for infectivity:** No information regarding the species of *Campylobacter* in the samples was available. Most human infections are caused by *C. jejuni* followed by *C. coli*; however, other species, including *C. upsaliensis*, *C. lari* and *C. fetus*, are occasionally associated with human infections. Identifying the species is therefore relevant to assessing the human health implications of the predicted concentrations.

Box C.3 Markov chain Monte Carlo and QMRA

- Markov chain Monte Carlo (MCMC) simulation is a random sampling technique that has been applied in many QMRA studies for the fitting of statistical models to microbial data sets (e.g. Teunis, Evers & Slob, 1999; Petterson, Teunis & Ashbolt, 2001; Petterson, Signor & Ashbolt, 2007; Teunis, Ogden & Strachan, 2008). In these studies, MCMC is used to explore the uncertainty in the parameter values for a statistical model or parametric distribution. MCMC has been used to define the credible intervals for the gamma distribution describing *Campylobacter* concentration in Fig. C.5 (a 19-parameter hierarchical model).
- When constructing a likelihood function with multiple parameters, the resulting function is complex and can be difficult to investigate. Within the framework of Bayesian inference, MCMC enables the posterior density to be explored by random sampling. Using a Markov chain algorithm (Gibbs Sampler or Metropolis-Hastings), rather than being a totally random series of independent samples (as in the case of Monte Carlo sampling), the chain is a random walk that proposes relatively small changes in location with each new iteration. The Markov chain approaches the stationary distribution for the posterior density. The final MCMC sample of all parameter values (after a burn-in phase) is assumed to be representative of the joint posterior distribution. For details on methodology, see Gilks, Richardson & Spiegelhalter (1996).

C4 Faecal indicators and quantitative prediction of pathogen concentration

Faecal indicator data sets are typically larger in size than pathogen data sets; in many situations, data on indicators are the only available microbiological data for characterization of local water quality. Not only is the measurement of indicators more straightforward (often leading to larger data sets), but the precision for enumeration is greater

than for pathogens (see section C2.1). Indicator data can therefore provide important information regarding the magnitude of and fluctuations in faecal contamination. Quantitative interpretation of the indicator data for estimating pathogen concentration requires an appreciation of the theoretical relationship between indicators and pathogens.

Even when pathogen data are available, it is generally beneficial to also include analysis of indicators alongside faecal source attribution data, such as via MST (e.g. Gourmelon et al., 2010; Schijven & de Roda Husman, 2011), to provide stronger evidence for the characterization of pathogen concentrations and likelihoods of human-infective strains for QMRA.

C4.1 Theoretical relationship between pathogens and FIO

Faecal indicators can originate from several sources; however, pathogens originate from only a subset of these. In addition, faecal indicators can exhibit very different environmental persistence from pathogens. The quantitative ratio between FIO and pathogens in a sample depends on:

- **Which** faecal indicator: Table C.5 summarizes the definitions of and health significance associated with the most commonly used FIO.
- From **whom** the faecal indicators are coming: indicators may be sourced from multiple hosts; however, human-infective pathogens are commonly found in only a subset of these.
- **Where** the sources are located, and **how** their faeces are deposited: consideration of the fate and transport of the indicators relative to the human pathogens will affect the relationship between their concentrations.

The expected relationship between pathogens and faecal indicators depends on how specifically faecal in origin the particular indicator is, how specifically human it is (i.e. MST value) and how environmentally persistent it is in comparison with human pathogens. Interpretation of indicator data therefore requires careful consideration of the faecal sources to the sample. For human sewage, the relationship between pathogens and indicators is much more consistent than for surface waters, as the **who** (humans) and the **where** (e.g. sewage treatment plant) are usually well defined. For surface waters, defining the **who** (humans, livestock, wild animals, etc.) and the **where** (large treatment plants, agricultural runoff, on-site systems, open defecation on land or from waterfowl, etc.), which can include numerous diffuse and/or point sources, requires a careful sanitary survey.

The expected relationship between pathogens and faecal indicators in a source will be driven by prevalence and population size of the contributing hosts. Even for a single, (relatively) well-defined source such as human sewage, the fluctuations in prevalence of pathogens will lead to considerable temporal variability in the pathogen to indicator ratio. Indicators such as *E. coli* are excreted by most people, whereas enteric pathogens are assumed to be excreted only by infected individuals, which may have a prevalence of much less than 1% (but up to 30% prevalence, e.g. for noroviruses in some populations). When the incidence of pathogen infections is low, sewage from a large population will provide a more consistent indicator to pathogen ratio compared with sewage from a small population or a single household. For example, consider a pathogen with a point prevalence of 0.1%. In sewage from a large community (100 000), around 100 people would be expected to be excreting the pathogen at any point in time. Although there will be some fluctuations over time with prevalence, the *E. coli* to pathogen ratio would be relatively constant. Conversely, for a single household (six persons) in low disease burden countries most of the time, no pathogens would be excreted; however, occasionally (approximately one in every 167 households), there would be one or more people excreting. During this event, pathogen concentration in the waste stream from that single household would be very high, with 1/6 people excreting. At the same time, however, because of a much higher prevalence, the indicator concentration would be expected to remain reasonably constant over time, even for a single household. As a result of this pattern, the pathogen to indicator ratio for the waste stream from a single household would not be very meaningful for risk assessment (fluctuation from zero to >1 for viruses); it would be more sensible to consider the likelihood of one or more infections in a single household and predict the pathogen concentration in the waste stream during that event. Conversely, for a large population with a small proportion of individuals excreting pathogens, the pathogen to indicator ratio would be much more constant and hence meaningful for risk assessment.

The expected relationship between indicators and pathogens will also depend on environmental or engineered barriers between the faecal sources and the sampling point. FIO and pathogens sometimes respond differently to environmental stressors and treatment barriers. For example, the reduction efficacy of sewage treatment barriers is different for bacteria, viruses and protozoa (and sometimes even for pathogen strains within these groups); hence, the thermotolerant coliform to rotavirus ratio would be expected to be different for raw sewage, primary treatment sewage and sewage following lagoon treatment. The application of a ratio needs to consider

the expected differences in behaviour between pathogens and indicators. In the same way, owing to differences in environmental persistence, the ratio of thermotolerant coliforms to rotavirus would be expected to change with time or distance downstream from a sewage source.

It is also important to consider the method of assay, such as culture versus PCR, which can, for example, yield highly different results for FIO compared with pathogens for sewage coming from different levels of treatment (Schoen, Soller & Ashbolt, 2011). These relationships in ratios are likely to be different when comparing pathogens with indicators within the same pathogen class, such as ratios of faecal indicator bacteria (thermotolerant coliforms or *E. coli*) to faecal bacterial pathogens (e.g. *Salmonella* spp.). Here, the ratios may not change as much, because both may respond similarly to environmental stressors and treatment barriers.

Table C.5 Summary of microorganisms commonly used as FIO or MST markers in environmental waters

FIO/MST marker	Definition	**Sources** Is the organism specifically of **faecal** origin?	**Hosts** Is the organism specifically of **human** origin?	**Environmental persistence** How **persistent** is the organism in comparison with human pathogens?	**Health significance** What is the health significance of finding the organism in surface waters?
Thermotolerant coliforms	All coliforms that can ferment lactose and produce acid (and gas in some jurisdictions) at 44.5 °C within 24 h; includes *Escherichia* and *Klebsiella* spp.	Tolerance of high temperature is suggestive of their faecal origin, yet some have environmental sources; in addition to faecal sources, environmental amplification is known (e.g. paper wastewater, tropical waters/ soils)	No	Persistence assumed similar to human bacterial pathogens; in general, much less persistent than viruses or protozoa	Presence in an environmental sample suggests recent faecal contamination; absence does not rule out the presence of more persistent pathogens
E. coli	Thermotolerant coliforms with the absence of urease and the presence of β-glucuronidase enzyme activity (although taxonomically up to 10% of environmental *E. coli* may not possess this enzyme, including pathogens such as *E. coli* O157:H7)	Mostly; excludes other species of bacteria within the thermotolerant group that are less likely to be of faecal origin	No	As for thermotolerant coliforms	As for thermotolerant coliforms
Faecal streptococci	A group of Gram-positive bacteria belonging to the genera *Enterococcus* (e.g. *E. faecalis*, *E. faecium*) and *Streptococcus* (e.g. *S. bovis*, *S. equines*)	Yes, but some environmental members	No	Tolerant of a wide range of environmental conditions; more tolerant than *E. coli*, but not as tolerant as many other non-bacterial pathogens	Demonstrated risk for recreational waters impacted by sewage, particularly by qPCR (Wade et al., 2010)

Table C.5 Summary of microorganisms commonly used as FIO or MST markers in environmental waters (contined)

FIO/MST marker	**Definition**	**Sources** Is the organism specifically of **faecal** origin?	**Hosts** Is the organism specifically of **human** origin?	**Environmental persistence** How **persistent** is the organism in comparison with human pathogens?	**Health significance** What is the health significance of finding the organism in surface waters?
Enterococcci/ intestinal enterococci	A subgroup of faecal streptococci including *Enterococcus faecalis* and *E. faecium*	Yes, but some environmental members (Badgley, Thomas & Harwood, 2010)	No; some species are more likely to be of human origin	As for faecal streptococci	Demonstrated for recreational waters impacted by sewage
Bacteriophages	A wide range of abundant viruses that infect bacteria	No	No	Diverse, but some may have properties assumed similar to those of human enteric viruses	None, as they are ubiquitous in the environment, but value as a treatment surrogate for enteric viruses
Somatic coliphages	Bacteriophages that infect *E. coli* via the cell wall and include spherical phages of the family Microviridae and various tailed phages in three other families	Most commonly of faecal origin, although environmental amplification is possible under some conditions	No, generally; some host bacteria provide greater human specificity	Assumed to be similar to human enteric viruses	As for bacteriophages
F-RNA coliphages	Bacteriophages with RNA genome that infect *E. coli* via the pili on cell surface	Yes	Groups 2 and 3 considered more human specific; some genotypes are largely human specific, groups 1 and 4 are not (Haramoto et al., 2012)	Generally more resistant than somatic coliphages, but less persistent than other markers (Jeanneau et al., 2012); relatively persistent and similar to enteric viruses	Main role as surrogates for treatment efficacy (Hata, Kitajima & Katayama, 2013); somewhat predictive of human health risk, especially recreational water
F-DNA phages	Bacteriophages with DNA genome that infect *E. coli* via the pili on cell surface	As for F-RNA phages	No	As for F-RNA phages	As for F-RNA phages
Bacteroidales	Order of strictly anaerobic bacteria containing some of the most numerous members of the gut microbiome; various members used to identify faecal sources by qPCR	Total Bacteroidales are not, but identified members are	Some, e.g. HF183 and HumH2 (Shanks et al., 2009; Haugland et al., 2010; Staley et al., 2012) and other members targeting other animal species (Toledo-Hernandez et al., 2013)	Less persistent than *E. coli* (Liang et al., 2012)	Provide confirmation of faecal input and to some degree source apportionment

Annexes

Table C.5 Summary of microorganisms commonly used as FIO or MST markers in environmental waters (contined)

FIO/MST marker	Definition	**Sources** Is the organism specifically of **faecal** origin?	**Hosts** Is the organism specifically of **human** origin?	**Environmental persistence** How **persistent** is the organism in comparison with human pathogens?	**Health significance** What is the health significance of finding the organism in surface waters?
Lachnospiraceae	Family of anaerobic gut bacteria, with members of the genus *Blautia* as faecal source indicators (McLellan et al., 2013)	Yes for members of genus *Blautia*	Yes for members of genus *Blautia*	Unknown	Provide confirmation of faecal input and to some degree source apportionment
Bacteroides fragilis bacteriophage	Bacteriophage that infects *B. fragilis*, with varying degree of human or source specificity (Diston, Ebdon & Taylor, 2012)	Yes	Phages to the host strain *B. fragilis* HSP40 are considered to be human specific, but phages to *B. fragilis* RYC2056 are more numerous and not human specific	Within the bacteriophages, *B. fragilis* is expected to be more persistent than other groups	Suggests human faecal contamination and likely presence of human enteric viruses
Human polyomaviruses	Group of human-specific enteric viruses excreted largely in urine (Polo et al., 2004; McQuaig et al., 2009)	Urine, less via faeces	Yes	Poor persistence, useful if recent human faecal contamination (McQuaig, Griffith & Harwood, 2012)	As for *B. fragilis* bacteriophages, but noting predominantly from urine, not faecal contamination
Sulfite-reducing clostridia	Gram-positive, spore-forming, non-motile, strictly anaerobic rods that reduce sulfite to hydrogen sulfide	No	No	Spores are highly resistant to environmental inactivation; more persistent than human pathogens	Uncertain; insufficient data; recreational water quality indicator in some settings
Clostridium perfringens	A sulfite-reducing anaerobic spore former that is exclusively faecal in origin	Yes	No	As for sulfite-reducing clostridia	Suggests faecal contamination that could be either recent or old; in the absence of *E. coli* or enterococci, *C. perfringens* is indicative of old/distant faecal contamination or disinfected wastewater
Hydrogen sulfide–producing bacteria	Diverse taxonomy of bacteria including sulfite-reducing clostridia and some (thermotolerant) coliforms	No	No	Some are more persistent than human pathogens	

Sources: Ashbolt, Grabow & Snozzi (2001); Griffith et al. (2009); Wade et al. (2010); Eichmiller, Hicks & Sadowsky (2013)

C4.2 Implementation of pathogen to indicator ratios in QMRA

Despite the uncertainty in interpreting pathogen concentrations from indicator results, there are several examples in the literature of using pathogen to indicator ratios for the purpose of quantifying reference pathogen concentrations in source material for QMRA. Data availability is the primary driver for using faecal indicators to quantify pathogen concentrations; even when available (which is rare), pathogen data sets are usually small, and the reliability of estimated concentrations is often low. When the epidemiology (patterns in prevalence and microbial sources) and environmental context (relative persistence and transport) are taken into consideration, faecal indicator data are of great value for QMRA (e.g. Petterson et al., 2015, 2016; Petterson, Stenström & Ottoson,

2016). In particular, the typically larger data sets of faecal indicators can provide invaluable context for risk assessment regarding the magnitude and fluctuations of faecal contamination.

Examples of the application of FIO to pathogen ratios from the literature include in drinking-water (Howard, Pedley & Tibatemwa, 2006; van Lieverloo, Blokker & Medema, 2007; Machdar et al., 2013), recreational waters (Craig, Fallowfield & Cromar, 2003; Soller et al., 2006) and wastewaters (Shuval, Lampert & Fattal, 1997; Mara et al., 2007; Bastos et al., 2008; Seidu et al., 2008). For these studies, the ratios applied and the data and assumptions on which they were based are summarized in Table C.6. In general, two different approaches have been applied in the development of the ratios:

- **Conservative limit:** Soller and coworkers (2006) selected a conservative enteric virus to coliphage ratio. From a review of eight studies from the literature, the authors noted that F-specific RNA coliphages were always present in numbers equal to or greater than the numbers of enteric viruses in ambient waters. The ratio of coliphage to viruses reported ranged from approximately 1:1 to approximately 1 000 000:1, with levels most commonly reported between 100:1 and 1000:1. A ratio of 1:1 was selected as a conservative value for the risk assessment. Similarly, Howard, Pedley & Tibatemwa (2006) used somatic coliphages as a conservative indicator (1:1) of rotavirus and sulfite-reducing clostridia (SRC) as a conservative indicator (1:1) of *Cryptosporidium* in drinking-water from various sources.
- **Average or expected value:** In the majority of studies, the ratio was developed based on a comparison between the average (specific statistical approaches varied or were not specified) concentration of the pathogens and the indicators from a particular data set that was assumed to then be representative for the risk assessment.

The approach taken in developing the ratio (including the size and representativeness of the data on which the ratio was based) is important for interpreting the risk assessment results. For example, the results from Soller et al. (2006) are conservative or represent an upper bound on the virus concentration and hence risk, whereas the risks assessed by Mara et al. (2007) (as argued in their publication) may be higher or lower than the actual risks, depending on the local fluctuations in pathogen concentration. Other factors will influence the interpretation of any risk estimates based on FIO, which is discussed in more detail by Howard, Pedley & Tibatemwa (2006) in section C4.3.

C4.3 Example: application of pathogen to indicator ratio in QMRA (Howard, Pedley & Tibatemwa, 2006)

Howard, Pedley & Tibatemwa (2006) explored the usefulness of the QMRA framework in developing contexts with limited data. Their study location was Kampala, Uganda, where water was drawn from the Inner Murchison Bay on Lake Victoria and passed to one of two treatment works: Gaba 1 and Gaba 2. In this example, no information was available on the pathogen concentration in raw water; all that was available was limited data on faecal indicator concentrations at the inflow to each plant. In order to use these data for QMRA, the authors made assumptions about the indicator to pathogen relationship.

Three reference pathogens were selected for the QMRA: pathogenic *E. coli*, *Cryptosporidium parvum* and rotavirus. FIO were used to estimate the pathogen concentration in raw waters at the Gaba 1 and Gaba 2 treatment plants: for pathogenic *E. coli*, 95% of thermotolerant coliforms were assumed to be *E. coli*, and 8% of *E. coli* were assumed to be pathogenic; for *Cryptosporidium parvum*, the concentration of SRC in the raw water was assumed to equal the *Cryptosporidium* concentration; and for rotavirus, the concentration of coliphage in raw water was assumed to equal the rotavirus concentration; this was argued to represent a worst-case scenario. The indicator data and the inferred pathogen concentrations are summarized in Table C.7.

Annexes

Table C.6 Examples of indicator to reference pathogen ratios applied in QMRA

Context	Indicator	Reference pathogen	Assumed ratio/approach for quantifying reference pathogen concentration	Cited reference
Drinking-water				
Howard, Pedley & Tibatemwa (2006) Various drinking-water sources (Uganda)	Thermotolerant coliforms	Pathogenic *E. coli*	95% thermotolerant coliforms assumed to be *E. coli* 8% of total *E. coli* assumed pathogenic Median number of *E. coli* or thermotolerant coliforms used as point estimate	WHO (1996) Haas, Rose & Gerba (1999)
	Somatic coliphages	Rotavirus	Somatic coliphage:rotavirus 1:1 worst case Median number of coliphage used as point estimate	Grabow (2001) cited for using coliphage as an index
	SRC	*Cryptosporidium*	SRC:*Cryptosporidium* 1:1 worst case. Median number of SRC used	Ashbolt, Grabow & Snozzi (2001) as reference to use SRC as an index
Machdar et al. (2013) Various water sources (Ghana) including household water storage, private taps, water sachets, water tankers and communal points	*E. coli*	*E. coli* O157:H7	1:0.08	Haas, Rose & Gerba (1999)
	E. coli	*Campylobacter*	1:0.66	1:0.9 in ozonated water and 1:0.66 average value cited to Smeets (2008)
	E. coli	Rotavirus	$1:5 \times 10^{-6}$	Wastewater (Mara et al., 2010)
	E. coli	*Cryptosporidium*	$1:10^{-6}$	Mara et al. (2010)
Wastewater reuse				
Shuval, Lampert & Fattal (1997) In wastewater for irrigation and on crops at consumption (Israel)	Faecal coliforms	Hepatitis A; rotavirus	$1:10^{-5}$	Various studies reviewed by Schwartzbrod (1995)
Mara et al. (2007) In wastewater of various qualities for irrigation and in soil (Mexico)[a]	*E. coli*	*Campylobacter*	$0.1–1:10^{-5}$	Data from waste stabilization ponds in north-eastern Brazil (Oragui et al., 1986)
	E. coli	Rotavirus	$0.1–1:10^{-5}$	Data from ponds in Kenya (Grimason et al., 1993)
	E. coli	*Cryptosporidium*	$0.01–0.1:10^{-5}$	
Recreational waters				
Craig, Fallowfield & Cromar (2003) Seawater (Australia)	Faecal coliforms	*Salmonella* spp.	Triangular distribution ($1:10^{-3}$; $1:10^{-4}$; $1:10^{-5}$)	Distribution selected to represent the range of values reported in the literature
Soller et al. (2006) Recreational waters (USA)	F-specific coliphages	Rotavirus	1:1 assumed to be conservative; reported values were variable	Eight studies

[a] These same ratios were also applied for wastewater-irrigated salad crops at harvest in Brazil (Bastos et al., 2008) and with faecal coliforms in irrigation waters of various qualities and on crops in Ghana (Seidu et al., 2008); and are used in the wastewater examples in the GWEG (WHO, 2006).

Table C.7 Faecal indicator data and inferred pathogen concentrations

Sampling location	No. of samples	Thermotolerant coliforms (median·100 mL^{-1})	*E. coli* (median·100 mL^{-1})	Assumed pathogenic *E. coli* concentration in raw water for QMRA (number·L^{-1})
Gaba 1, 2002	36	–	42 (<1–TNTC)	32
Gaba 2, 2002	19	–	20 (2–120)	18
Gaba 1, 1994–1998	200	15	–	11.5
Gaba 2, 1994–1998	200	10	–	7.5
Distribution system, 1998	713	0.23 (<1–17)	–	0.18
Distribution system, 1999	913	0.13 (<1–30)	–	0.10
Household-stored water, 1999	97	3 (<1–TNTC)	–	2.3
Protected springs, 1998–1999	609	14 (<1–TNTC)	–	10.6
	No. of samples	**SRC (median·100 mL^{-1})**		**Assumed *Cryptosporidium parvum* concentration in raw water for QMRA (number·L^{-1})**
Gaba 1, 2002 and 2003	36	3 (<1–TNTC)		30
Gaba 2, 2002 and 2003	19	5 (1–5)		50 (210[a])
	No. of samples	**Somatic coliphage (median·mL^{-1})**		**Assumed rotavirus concentration in raw water for QMRA (number·L^{-1})**
Gaba 1, 2003	36	<1		1 000
Gaba 2, 2003	19	1		900[b]

[a] On one sampling occasion, a concentration of 21 CFU·mL^{-1} was reported in post-treatment water; this was included separately in the QMRA.

[b] Estimated based on non-detection in 1 mL and default concentration of 0.9 mL^{-1}.

Source: adapted from Howard, Pedley & Tibatemwa (2006)

The risks from pathogenic *E. coli* associated with consumption of water leaving the drinking-water treatment plants (assuming 5 $\log_{10}$ reduction for treatment performance), water from the distribution system, household water and spring water were compared with a reference level of tolerable risk (10^{-6} DALY) and with each other. The following conclusions were drawn:

- Disease burden associated with the water supplies varied between sources. Water leaving the treatment plants was comparable with the reference risk level; water in the distribution system was around 2–4 $\log_{10}$ above the reference level; and water from household storage and the protected springs was 4–6 $\log_{10}$ above the reference level.
- The authors concluded that risks associated with distributed water were higher than risks associated with water leaving the treatment plant, identifying the need for improved operation and maintenance of the distribution system.
- Risks associated with spring water and stored household water were higher than those associated with piped supplies, highlighting the need to increase access to piped supplies.

For *Cryptosporidium*, the risk from water leaving each treatment plant was compared with the reference level of tolerable risk and was calculated to be around 1 $\log_{10}$ higher than the reference level at Gaba 1 and 2 $\log_{10}$ higher than the reference level at Gaba 2. The risk was also calculated for a failure event at Gaba 2, during which there was a significant 4 $\log_{10}$ increase in comparison with the average. Similarly, for rotavirus, the risk was calculated to be around 3.5 $\log_{10}$ higher than the reference level from both plants.

For a full discussion of the model inputs and uncertainty associated with the assumed pathogen concentrations, the reader is referred to the original publication. It is clear, however, that the validity of the QMRA outcomes depends on

the inferences made regarding the relationship between indicators and pathogens in order to predict the pathogen concentrations in Table C.7. In particular, for the *E. coli* analysis, the validity of the conclusions associated with the comparison of risks between water sources rests on the assumption that the health significance associated with the presence of *E. coli* and thermotolerant coliforms in untreated water, distributed water, household water and spring water is the same. By implication, this assumes that the sources of indicators to untreated water, distributed water, household water and spring water have identical indicator to pathogen ratios.

In this study, only faecal indicator data were available; although the assumption of "same health significance" introduced a constraint to the interpretation and a source of uncertainty, it did allow for the water sources to be quantitatively compared from a disease burden perspective. The results could be used to support decisions regarding infrastructure development and investment, an important benefit and application of the indicator monitoring data. If there was compelling evidence to support different faecal sources between water supplies, whether or not the "same source assumption" could have biased the outcomes of the risk assessment could be further explored through scenario analysis, assuming different faecal sources to each water supply.

C5 Quantifying pathogen reduction across barriers: empirical approach

C5.1 Data considerations

When adopting an empirical approach, observations are collected before and after the treatment barrier and statistically analysed to provide a quantification of the treatment efficacy. Regarding the observational data, the factors described below need to be considered.

C5.1.1 Surrogates

Evaluation of treatment efficacy can rarely be undertaken using observations of human pathogens. This is because pathogen concentrations are typically too variable and low (particularly after a barrier) to allow for enumeration and effective evaluation of barrier efficacy. Surrogates (native and spiked) are therefore used to evaluate removal performance. Where mechanisms of removal are not isolated, the surrogate is assumed to mimic the behaviour of the pathogen in all aspects; nevertheless, the primary mechanism of removal (e.g. filtration, inactivation) will influence the selection of a suitable surrogate. Although, historically, faecal coliform bacteria have been most commonly used to describe pathogen reduction across a barrier, this group of organisms cannot be expected to behave in the same way as the diverse range of pathogens in terms of physical size of the microbe and persistence. Microbial size, persistence and surface properties are typical determinants of surrogate selection, and therefore different surrogates will be selected to represent bacterial, viral and protozoan pathogens.

Cryptosporidium removal has been modelled using aerobic spores (Medema et al., 2003; Jaidi et al., 2009); spores of SRC (Teunis et al., 1997; Havelaar et al., 2000; Medema et al., 2003; Howard, Pedley & Tibatemwa, 2006; Oesterholt et al., 2007; Carlander, Schönning & Stenström, 2009); *Clostridium perfringens* (Ottoson & Stenström, 2003); particles (1–15 µg) (Westrell et al., 2003); and PRD1 bacteriophage (Medema et al., 2003). Using the same data sets, some studies also modelled *Giardia* with SRC (Teunis et al., 1997; Oesterholt et al., 2007; Carlander, Schönning & Stenström, 2009) and *Clostridium perfringens* (Ottoson & Stenström, 2003). Physical removal of viruses has been modelled using somatic coliphages (Ottoson & Stenström, 2003; Howard, Pedley & Tibatemwa, 2006; Oesterholt et al., 2007; Carlander, Schönning & Stenström, 2009) and F-specific RNA coliphages (Oesterholt et al., 2007). *Escherichia coli* was used to model the removal of *Salmonella* (Carlander et al., 2009), *Campylobacter* (Oesterholt et al., 2007) and pathogenic *E. coli* (Howard, Pedley & Tibatemwa, 2006) across water treatment. Enterococci were used to model removal of *Campylobacter* (Ottoson & Stenström, 2003; Oesterholt et al., 2007) in greywater and household water treatment and pathogens through superficial aquifers (Roser et al., 2002). It is important, however, to understand the limitations associated with use of these surrogates (Payment & Locas, 2011).

In wastewater reuse studies, non-microbial surrogates have also been used to estimate the magnitude of wastewater or faecal contamination. Shuval, Lampert & Fattal (1997) relied on the volume of wastewater attached to irrigated lettuce in order to estimate microbial contamination of wastewater-irrigated crops. This approach was extended to broccoli and cabbage by Hamilton and co-workers (2006). Faecal sterols were used as a surrogate for faecal contamination of urine (Höglund, Stenström & Ashbolt, 2002) and greywater (Ottoson & Stenström, 2003).

C5.1.2 Sampling regime

When using an empirical approach, observed reductions are representative only for the specific experimental conditions. Therefore, understanding the conditions under which the samples were collected is critical to interpreting the results. Three particular factors require consideration:

1) **Variability in inflow and outflow concentrations:** For real barriers (pilot- and full-scale), the concentrations of pathogens and native surrogates are constantly fluctuating in the samples before and after the barrier. These fluctuations, particularly at the time of sampling, will influence the interpretation of the results. *Note:* Comparison of a single sample before and after treatment can imply an increase in concentration; however, this phenomenon is often the result of random fluctuations in concentration (e.g. a low random sample before, and a high random sample after) that can occur even when the overall concentration of organisms is reduced across the barrier.
2) **Variability in barrier performance:** Even for barriers operating under stable (steady-state) conditions, the treatment performance may be variable. The time scale of the experimental data is therefore important. Studies may collect many samples before and after treatment over a short period (e.g. many samples collected on the same day, before and after treatment), or samples may be collected over a longer period (e.g. monthly samples before and after treatment over a year).
3) **Were samples paired before and after the barrier?** Samples are often collected on the same day and at a similar time. Whether or not the samples are paired (represent the same water/media before and after the barrier) depends on the residence time and dynamics of the individual barrier. When samples are paired, the uncertainty in removal performance can be reduced.

C5.1.3 Challenge testing

Challenge testing to characterize treatment efficacy (where the inflow to the treatment barrier is spiked with a high concentration of microbial surrogates) is often undertaken in order to overcome the problems associated with relying on native surrogates. Guiding principles for challenge testing of household treatment devices to characterize efficacy are given in the WHO document on evaluating household water treatment options (WHO, 2011); and examples of pilot-scale (Hendricks et al., 2005) and full-scale challenge testing of water treatment barriers have been published in the literature (e.g. Davies et al., 2008).

C5.2 Statistical analysis

The following factors need to be considered in the selection of a suitable statistical approach or model for quantifying barrier performance. The approach will depend on the value to be inferred, which may include:

- the expected value of the barrier performance;
- the minimum value of the barrier performance;
- the range of expected performance; and
- the probability distribution describing the variability in removal performance.

Do the data represent the original observations or reported concentrations? Monitoring data are typically reported as concentrations. These concentrations are an inference based on the laboratory observations (see section C2).

What are the characteristics of the data? These characteristics (qualitative, discrete, categorical or continuous) will influence the selected modelling approach, as the selected distributions must be appropriate to the type of data (see section C3).

Four statistical approaches have been used within QMRA to quantify microbial reduction efficacy. Each approach has different assumptions regarding the variability in microorganism concentration (normal, lognormal or another alternative distribution) and variability in treatment (is it constant or variable?) and the interpretation of microbial data sets. The four statistical approaches are described below:

1) **Single point estimates** to describe reduction efficacy for a particular process or treatment train; sometimes referred to as the decimal elimination capacity. Barrier efficiency studies in the literature typically report a point estimate of reduction that represents the $\log_{10}$ difference in mean concentrations before and after the barrier (e.g. for drinking-water treatment processes: Payment & Franco, 1993; Hendricks et al., 2005; Hijnen et al., 2007; and wastewater treatment: Rose et al., 1996; Scott et al., 2003). Using this approach, the method of averaging the microbial counts can influence the quantified reduction; this is illustrated in section C5.3.

2) **Single point estimates by sampling occasion** (paired data): A range of point estimates that describe the $\log_{10}$ reduction in average concentration between sampling occasions (data paired before and after treatment pairs) – for example, greywater treatment (Ottoson & Stenström, 2003) and urban stormwater barriers (Davies et al., 2008).
3) **A ratio between the distributions** of inflow and the outflow concentration – for example, the ratio calculated by sampling from the Poisson lognormal distributions fitted to counts before and after treatment (Masago et al., 2004); and the likelihood function describing the parameters of the distribution of the ratio between two gamma variables fitted to virus counts before and after several different treatment barriers (Teunis et al., 2009).
4) **Statistical models** that describe microbial passage as a binomial (two-outcome) process (Teunis, Evers & Slob, 1999; Medema et al., 2003; Teunis et al., 2009).

The statistical approach employed will have an impact on the inferred $\log_{10}$ reduction of barrier efficacy. In general, as more factors are accounted for in the statistical analysis, the mathematical complexity increases. The more detailed statistical approaches (points 3 and 4 above) have been explored only in the context of drinking-water QMRA. Section C5.4 illustrates the impact that the analysis approach can have on the quantified $\log_{10}$ reduction. The influence of the analysis approach is expected to be greater for small data sets.

C5.2.1 Accounting for differences between the surrogate and the pathogen

When a surrogate is used to quantify pathogen reduction efficacy, there is uncertainty regarding how well that surrogate is able to represent the pathogen of interest. In most cases, quantitative information relating the pathogen and surrogate performances is not available. The predominant approach within QMRA is to simply assume that the surrogate behaves identically to the pathogen (which ignores a potentially important source of uncertainty). However, two exceptions have been published in the literature: Jaidi et al. (2009) attempted to quantify the relationship between oocyst and aerobic spore removal based on literature data; and Medema et al. (2003) used MS2 and PRD1 bacteriophage to evaluate *Cryptosporidium* oocyst removal by soil passage, applying a 1.5- to 3-fold increase to the estimated phage removal based on the results of Schijven et al. (1999, 2003).

C5.3 Example: quantifying $\log_{10}$ reduction from microbial counts

Microbial counts analysed from samples collected before and after a drinking-water treatment process are included in Table C.8.

Table C.8 Counts before and after a drinking-water treatment barrier

Before		After	
n	***V*** **(mL)**	***K***	***W*** **(mL)**
10	1.0	4	20
6	1.0	2	10
12	2.0	0	10

Three potential approaches for quantifying $\log_{10}$ reduction based on the data in Table C.8 are as follows:

1) Take ratios of concentrations before and after, and average:

$$R_1 = \frac{1}{3}\sum_{i=1}^{3} \frac{k_i/W_i}{n_i/V_i} = 0.017\ 78\ (\log_{10} \text{ reduction} = 1.75)$$

2) Average concentrations before and after, and use their ratio:

$$R_2 = \frac{\frac{1}{3}\sum_{i=1}^{3} k_i/W_i}{\frac{1}{3}\sum_{i=1}^{3} n_i/V_i} = 0.018\ 18\ (\log_{10} \text{ reduction} = 1.74)$$

3) Add counts and volumes before and after, and use their ratios:

$$R_3 = \frac{\sum_{i=1}^{3} k_i/\sum_{i=1}^{3} W_i}{\sum_{i=1}^{3} n_i/\sum_{i=1}^{3} V_i} = 0.021\ 43\ (\log_{10} \text{ reduction} = 1.67)$$

The same data are analysed; however, different estimates of the $\log_{10}$ reduction are obtained. Why are they different? R_1 assigns equal weights to ratios of concentration; R_2 assigns equal weights to concentrations; and R_3 weights the concentrations by the counted numbers. In the case of Poisson-distributed counts, R_3 is the best (maximum likelihood) estimator of the ratio of concentrations after/before the barrier. Does the difference matter? That depends on the purpose of the risk assessment and the required level of precision. The importance of the difference can be explored through sensitivity analysis during risk characterization (Chapter 8). Of most importance is to recognize that the assumptions behind the data analysis approach will influence the quantitative prediction of barrier efficiency.

C5.4 Example: reduction of *Giardia* across conventional drinking-water treatment

Monthly water samples were collected at the inflow and the outflow of a drinking-water treatment plant over the period of 1 year and were analysed for *Giardia*. The counts are summarized in Table C.9.

Table C.9 *Giardia* counts enumerated from inflow and outflow of a conventional drinking-water treatment barrier, with estimated concentrations and $\log_{10}$ removal rates

Sampling day	Inflow			Outflow			Reduction
	Count	Volume	Estimated concentration (cysts·L^{-1})	Count	Volume	Estimated concentration (cysts·L^{-1})	Estimated removal ($\log_{10}$ reduction)
1	3	10	0.3	0	100	0	>0.52
2	6	10	0.6	14	100	0.14	0.63
3	2	10	0.2	7	100	0.07	0.46
4	2	10	0.2	3	100	0.03	0.82
5	0	10	0	0	100	0	–
6	0	10	0	0	100	0	–
7	1	10	0.1	6	100	0.06	0.22
8	3	10	0.3	0	100	0	>0.52
9	6	10	0.6	0	50	0	>0.22
10	2	10	0.2	0	50	0	>0.70
11	30	10	3	18	100	0.18	1.22
12	0	5	0	2	100	0.02	–
Average of all days			0.46			0.042	1.04

In this example, four different analytical approaches were compared. The approaches and underlying assumptions are summarized in Table C.10 and discussed below:

Table C.10 Overview of statistical approaches for analysis of microbial reduction data

Analytical approach	Assumptions	What is not accounted for in the statistical approach?	Expected value of $\log_{10}$ reduction	Estimated variability in $\log_{10}$ reduction
Overall average performance	• The method recovery is constant between samples; and before and after treatment • Inflow and outflow *Giardia* concentration follows a normal distribution • Inflow and outflow concentrations are not correlated by sampling day	• Variability in the recovery of the analysis method • Sampling variability in *Giardia* counts • Correlation by sampling day before and after treatment • Variability in treatment performance	1.04	NA
Performance paired by sampling occasion	• The method recovery is constant • *Giardia* concentration is constant over a single sampling period, and therefore water sampled at the outflow can be considered to be exactly the same water that was sampled at the inflow	• Variability in the recovery of the analysis method • Sampling variability in *Giardia* cysts • Short-term fluctuations in *Giardia* concentration	1.04	0.22–1.22
Ratio between distributions of concentrations before and after treatment	• The method recovery is constant • *Giardia* cysts are randomly distributed in the water column and therefore can be described by the Poisson distribution • Variability in *Giardia* concentration can be described by a gamma distribution • Correlation between the inflow and outflow concentrations was either zero or complete	• Variability in the recovery of the analysis method	No correlation: 1.33 Complete correlation: 1.34	0.7–4.51 0.92–2.76
Binomial models	• The method recovery is constant • *Giardia* cysts are randomly distributed in the water column and therefore can be described by the Poisson distribution • Variability in *Giardia* concentration can be described by a gamma distribution • Pairing of samples by sampling day • Treatment reduction is variable, and the probability of cyst passage can be described by a beta distribution	• Variability in the recovery of the analytical method	0.92	0.37–3.3

NA: not applicable

1) **Overall average performance.** Barrier performance data are often analysed simply by comparing the average inflow and outflow concentrations. In this case: $\log_{10}$ (0.46) – $\log_{10}$ (0.042) = 1.04.

2) **Performance paired by sampling occasion.** The treatment performance was calculated for each individual sampling day. In this case: $\log_{10}$ reduction would be reported as 1.04 (0.22 – 1.22). Using method 2, some indication of the variability in the treatment performance between days is given; however, this is limited by the assumption of direct correlation between paired samples and the influence of zero counts. When zero organisms are counted at the outflow, the reduction can only be estimated to be "at least" the removal of the inflow concentration, and the upper limit of the reduction efficacy on that day cannot be estimated. In addition, when zero organisms are counted at the inflow and the outflow, no information is available regarding the reduction on that particular day.

3) **Ratio between distributions of concentrations before and after treatment.** The treatment performance can be characterized by fitting a distribution to the concentration both before and after treatment and then using Monte Carlo simulation to calculate the distribution of the ratio between the two distributions. The distribution for the ratio is then an estimator of the barrier efficiency.

   Fig. C.6 illustrates the gamma distributions fitted to the counts from the inflow and the outflow. The ratios between these two distributions, assuming both no correlation and complete correlation (random samples of inflow and outflow distributions were sorted, forcing complete rank correlation), are reported in Table C.10. The expected values of both analyses are similar (–1.33 versus –1.34); however, the variability (and uncertainty, as they are not separated in this model) in the distribution for *Giardia* removal is greatly reduced in the correlated model in comparison with the independent model. The true result would be expected to lie somewhere between these two unrealistic extremes.

4) **Binomial models:** Models have been presented in the literature that describe organism passage through treatment as a binomial process (Teunis, Evers & Slob, 1999), where each microorganism faces one of two possible outcomes: passage or removal. Mean microorganism concentration in the inflow (μ) is assumed to follow a gamma distribution. The maximum likelihood beta distribution for removal is illustrated in Fig. C.7.

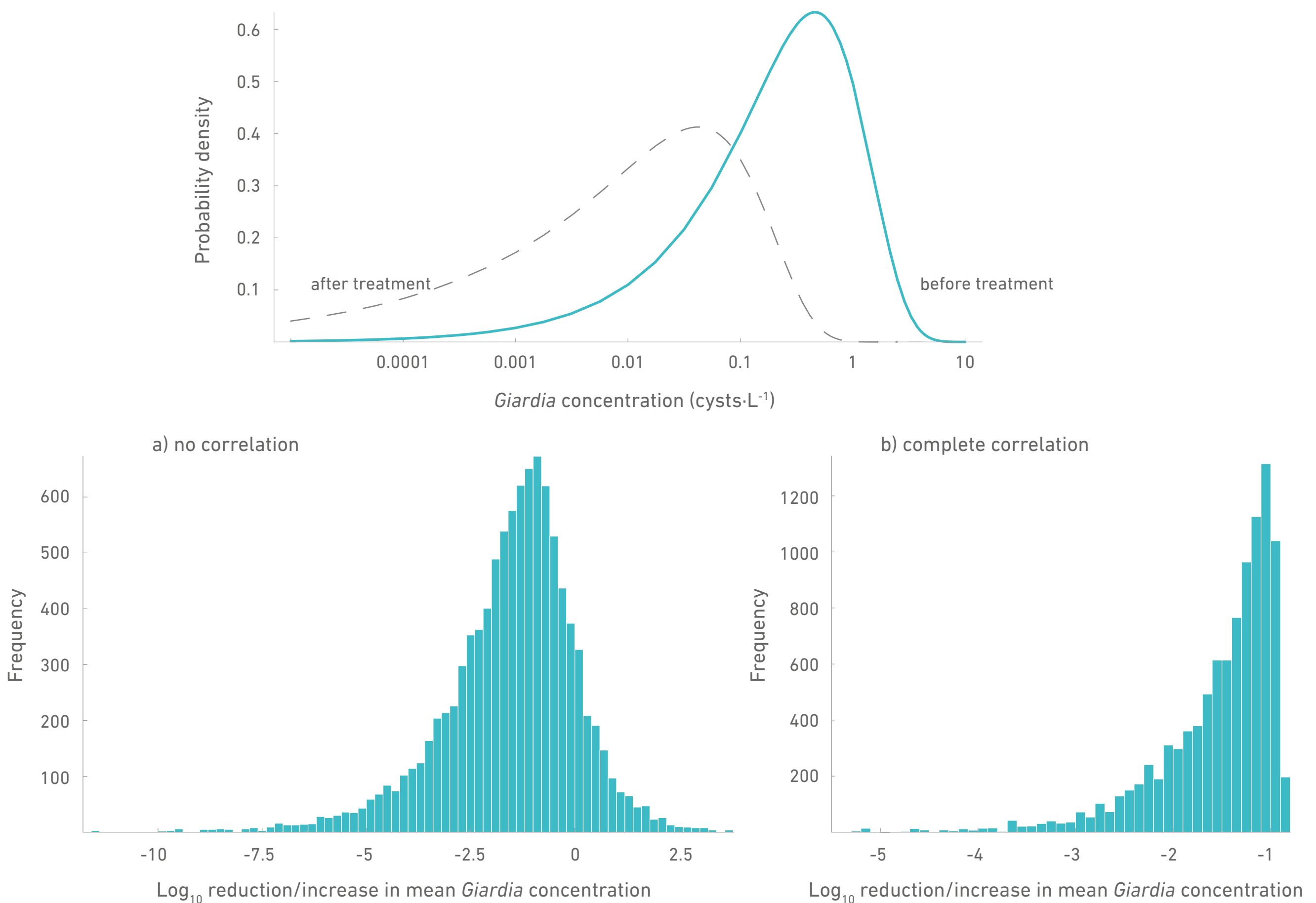


Fig. C.6 Gamma distribution fitted to *Giardia* counts at the inflow and outflow of the treatment barrier (top). Histogram of Monte Carlo sample of ratio between the two distributions with a) no correlation and b) complete correlation.

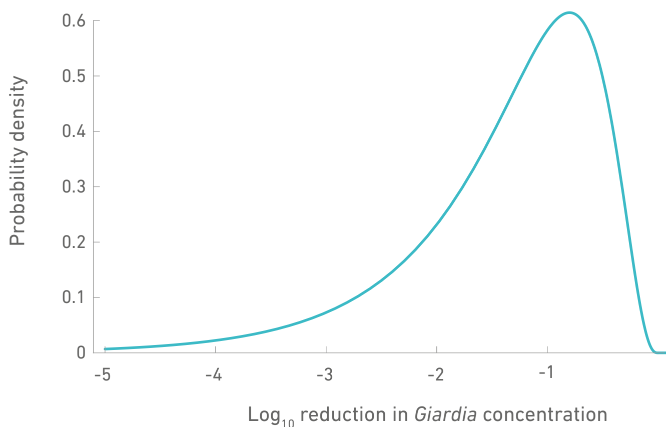


Fig. C.7 Maximum likelihood beta distribution describing $\log_{10}$ reduction across the treatment barrier, applying the beta binomial model with paired counts

In the analyses presented, the same data were interpreted to yield different quantitative inputs for the QMRA. Table C.10 shows the assumptions that are accounted for in each of the methods. The higher the number of factors accounted for in the analysis, the greater the mathematical complexity. Selection of the most appropriate approach will depend on the purpose of the QMRA, as defined in problem formulation, and the required level of detail in the analysis for the QMRA. When applying a tiered approach (Chapter 5), a conservative point estimate can be applied first, with more detailed analysis to follow, if required.

C5.5 Example: reduction across a full-scale urban wetland

A monitoring programme was undertaken to evaluate the microbial removal performance of a full-scale urban wetland in Melbourne, Australia (Petterson et al., 2016). The wetland was one component of an overall treatment system (illustrated in Fig. C.8) designed to collect and reuse urban stormwater for public space irrigation.

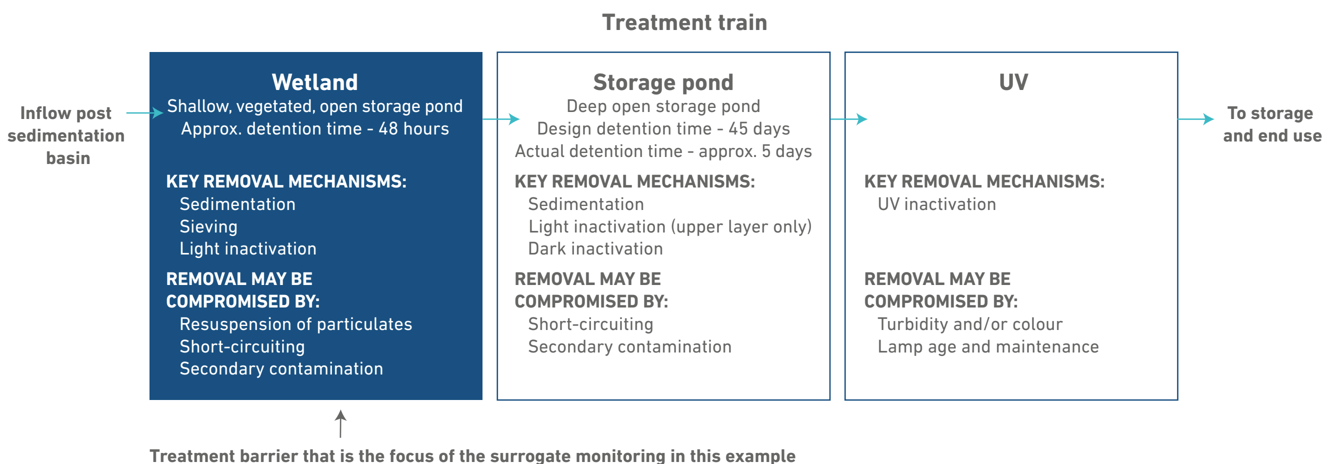


Fig. C.8 Overall treatment train for stormwater reuse in Melbourne, Australia (adapted from Mitchell et al., 2008)

Water samples were collected at the inflow and the outflow of the wetland and analysed for three microbial surrogate organisms: *E. coli*, intended to represent bacteria, somatic coliphage, to represent viruses, and *Clostridium perfringens*, to represent protozoa. Samples were collected under background (baseflow) conditions and during rainfall events using a flow-triggered auto-sampler. Results are summarized in Table C.11. The $\log_{10}$ reductions were calculated as the $\log_{10}$ difference in the geometric mean before and after the treatment barrier.

The results indicate that higher removals were apparent for all organisms under event conditions in comparison with background conditions. For *C. perfringens* and somatic coliphages, the concentrations at the outflow under event and background conditions were similar; however, the event inflow concentrations were higher. For *E. coli*, the inflow concentration was approximately 2 $\log_{10}$ higher under event conditions; however, the outflow concentration

was only around 1 log_{10} higher. These results imply that the wetland was effectively providing a buffer for the retention of microorganisms under event conditions; that is, whereas the geometric mean and range of inflow concentrations were much higher under event conditions, the geometric mean outflow concentration remained relatively steady.

Table C.11 Removal of native surrogate organisms by an urban wetland

	Event		Background		All	
	Inflow[a]	Outflow[a]	Inflow[b]	Outflow[b]	Inflow	Outflow
***E. coli* (MPN·100 mL^{-1})**						
Geometric mean Range [min–max]	11 000 [63–87 000]	581 [26–7 700]	202 [10–7 300]	66 [10–7 300]	2 525 [10–87 000]	375 [10–7 700]
Sample size [non-det]	54 [0]	71 [0]	31 [0]	18 [0]	85 [0]	89 [0]
Log_{10} removal	1.3		0.49		0.83	
***C. perfringens* (CFU·100 mL^{-1})**						
Geometric mean Range [min–max]	1 060 [1–9 000]	30 [0–670]	50 [1–1 000]	17 [0–1 500]	426 [0–9 000]	25 [0–1 500]
Sample size [non-det]	54 [0]	71 [23]	23 [0]	14 [5]	77 [0]	85 [28]
Log_{10} removal	1.6		0.46		1.2	
Coliphage (PFU·100 mL^{-1})						
Geometric mean Range [min–max]	360 [29–6 300]	84 [8–2 000]	49 [3–490]	74 [4–2 300]	199 [3–6 300]	71 [4–2 300]
Sample size [non-det]	54 [0]	71 [0]	23 [0]	14 [0]	77 [0]	85 [0]
Log_{10} removal	0.63		−0.18		0.45	

max: maximum; min: minimum; non-det: number of non-detects

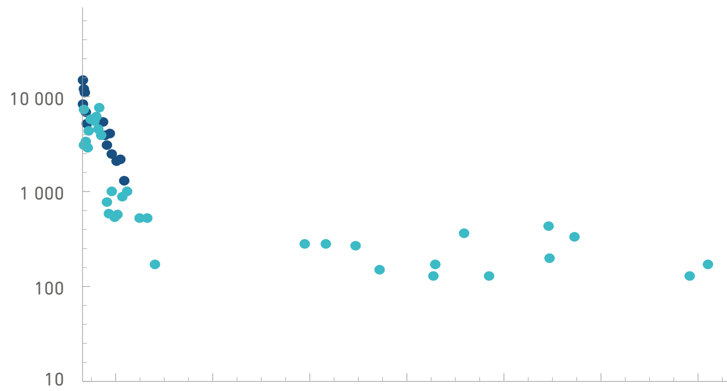
[a] Combined small event 1, small event 3, medium event 1, medium event 2 and large event 1.

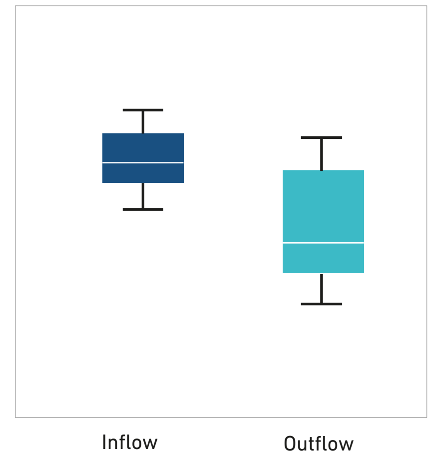
[b] Remaining data set.

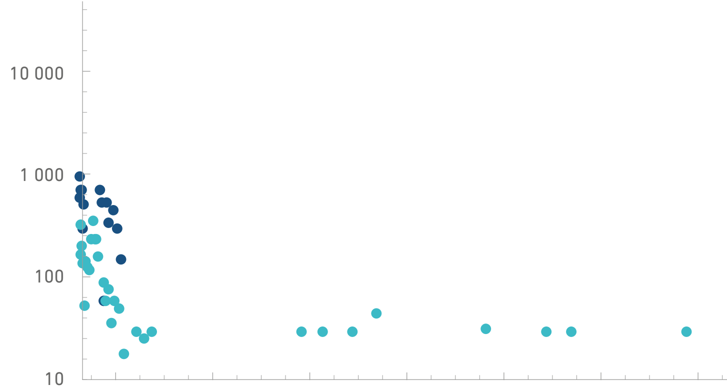
Source: Petterson et al. (2016)

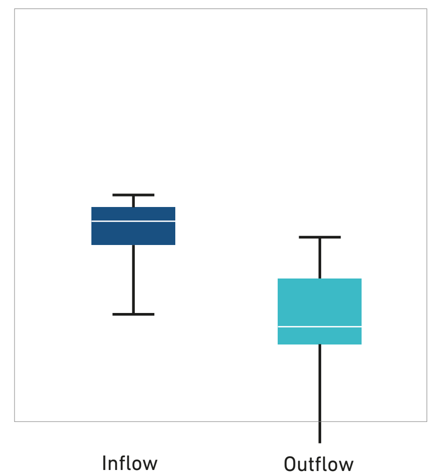
The simple comparison between the inflow and outflow concentrations of the wetland does not give the full picture of the microbial removal performance, as the system is not operating in steady state. Rather, the treatment process is dynamic and responding to pulses of flow and microbial loading due to rainfall events. Compare the estimated log_{10} reductions with the time series plot of a single event in Fig. C.9. For *E. coli*, the peak concentrations at the inflow and outflow occurred at the start of the event, with concentrations declining over time. These results suggest that an initial peak of *E. coli* was released from the wetland in response to the rainfall event. This peak could have been caused by one or a combination of two factors: 1) penetration of runoff through the wetland during high-flow conditions, enabling an initial plume of runoff to reach the outlet with only minimal reduction (peak inflow = 15 000 MPN·100 mL^{-1}; peak outflow = 7700 MPN·100 mL^{-1}, indicating a peak reduction of 0.29 log_{10}); in this case, *E. coli* would represent material from the catchment mobilized during the runoff event; or 2) resuspension of residual material in the wetland, stored from previous runoff events or generated by waterfowl.

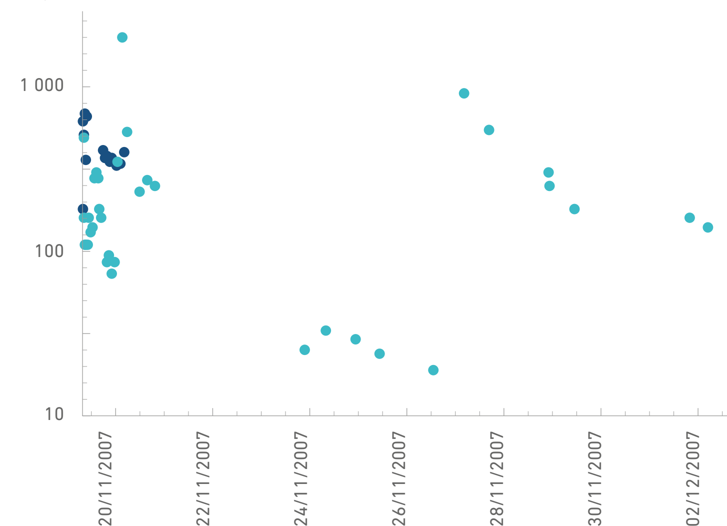
Consideration of the results for *C. perfringens* and somatic coliphages indicates a similar pattern, although it is less marked than for *E. coli*. In fact, for the coliphages, the initial peak exceeded the measured inflow concentration, and a second peak was apparent around 6 days after the initial rainfall event. It is not possible to determine if these results are a product of the inherent variability in the concentration or are a process-driven response to the inflows to the system.











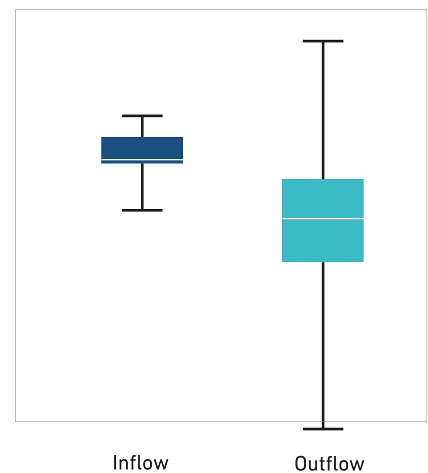


Fig. C.9 Surrogate monitoring results following a rainfall event across an urban wetland (dark blue dots represent wetland inlet, light blue dots represent wetland outlet; middle line represents median, box represents upper 75th and lower 25th percentiles, whiskers represent minimum and maximum) (Petterson et al., 2016)

A simple overall concentration comparison would ignore the potential risks associated with short-term penetration of pathogen peaks following runoff events. The risk assessment must give consideration to the dynamic nature of the wetland system as a removal barrier. The resulting risk characterization would not represent the fluctuations in risk associated with the system. Using the reported data, the short-term risks immediately following a rainfall event could be characterized separately from the baseline risk. Although this would be conservative, it would provide an indication of the fluctuations. Alternatively, a process-driven mechanistic approach to modelling the indicator data may reduce the uncertainty in characterizing the performance of the barrier in response to pathogen loading events.

C6 Quantifying pathogen reduction across barriers: mechanistic approach

Mechanistic, process-driven models are being increasingly used to describe the fate and transport of pathogens across barriers. Rather than simply comparing data sets before and after the barrier, process-based models consider factors driving the pathogen removal or inactivation and, moreover, seek to quantify those factors for the individual system. The two most common applications of a mechanistic approach in QMRA have been for hydrologic modelling and modelling pathogen inactivation.

C6.1 Hydrologic models

Hydrology is the study of the movement, distribution and quality of water. For water-related QMRA, the vehicle of disease transmission from source to exposure is water. Hydrologic models can thus be of great value for quantifying the fate and transport of pathogens under a diverse range of water conditions, including:

- rainfall and runoff models of water catchments, for describing the frequency and magnitude of overland flow events that mobilize pathogens to surface waters (Signor et al., 2005);
- hydrodynamic models of rivers, for describing the dispersion (dilution) and advection (transport time) of pollutants between an input and a downstream point of exposure or offtake location;
- hydrodynamic models of reservoirs, for describing the variability (in particular, potential for short-circuiting) in residence time under the range of inflow (rainfall and runoff) and stratification conditions (Brookes et al., 2004);
- hydraulic models of water and wastewater treatment facilities, for describing variable residence time within sedimentation basins, storage ponds and contact chambers;
- hydraulic models of distribution systems, for describing the impact of low-pressure ingress events and the fate and transport of introduced pathogens (Teunis et al., 2010b); and
- hydrogeologic models, for describing the fate and transport of waterborne pathogens in the subsurface (Schijven & Hassanizadeh, 2000; Schijven et al., 2003, 2006).

Drawing on the tools available for modelling of water movement and combining with an understanding of microbiological persistence provide a depth of understanding to the drivers of microbial risks that cannot be ascertained from the microbiological monitoring data alone. In particular, understanding the limits of water transport provides a boundary around the assessment of microbial risks (e.g. minimum transport times, short-circuiting, potential impact or travel distance of a contamination event).

C6.2 Pathogen inactivation

Pathogen inactivation outside the host, either as a result of natural environmental stressors or in the presence of a disinfectant, is an important barrier for the protection of public health. The implications of hydrologic models in terms of travel times from source to exposure can be interpreted for human health risk only with an appreciation of microbial persistence. Inactivation models are used to describe pathogen persistence in the environment (e.g. soil, water, crops), during thermal treatment (e.g. composting, cooking) and during disinfection (e.g. chlorine based, UV, ozone). For a given set of conditions, the rate of microbial inactivation varies between pathogen groups and between individual pathogen types.

Microbial persistence within QMRA is commonly modelled as a first-order process related to time: Surviving Fraction = $e^{-k \cdot t}$. Inactivation rates may be directly reported as k values (inactivation rate of a $\log_{10}$ scale) or as T_{90} (time for 1 $\log_{10}$ reduction) or T_{99} (time for 2 $\log_{10}$ reduction) values. For environmental inactivation, an overview of factors influencing pathogen survival and expected survival times (and, in some cases, expected decay rates; Table 3.8 in WHO, 2006) in various environmental media, including water, soil and faeces, and on food crops is provided in Chapter 3 of the GWEG (WHO, 2006).

For disinfection, inactivation is typically modelled as a first-order process related to the concentration of disinfectant × time (ct value) for disinfection: Surviving Fraction = $e^{-kc \cdot t}$. Indicative required ct_{99} (ct for 2 log_{10} inactivation) values for different pathogens and disinfectants are summarized in the GDWQ (Table 7.7; WHO, in preparation) and supporting documents (Sobsey, 1989; LeChevallier & Au, 2004). There is, however, considerable variability in the reported ranges for the value of microbial inactivation rates. Using these assumptions of first-order decay and a known decay rate, the log_{10} reduction can be estimated for a given set of barrier conditions. In particular, if the residence time of the contactor and disinfectant dosing is well characterized, then the expected reduction can be estimated for the specific contactor. Under these mechanistic assumptions and applying pathogen-specific inactivation data, a screening-level modelling approach to support the QMRA of drinking-water systems has been presented (Petterson & Stenström, 2015).

In selecting an inactivation rate from the literature, the following factors need to be considered:

- **Model organism:** Studies are often conducted on surrogate organisms, rather than on pathogens. What was the model organism, and what is known about how representative it is for the pathogen of interest?
- **Laboratory strains of organisms:** Laboratory experiments require a high concentration of target organisms to seed. These organisms are grown from laboratory strains. How well do these organisms represent the resilience of environmental strains?
- **Microbial methods:** For pathogens and microbial surrogates, all the methodological issues addressed in sections C1 and C2 are also relevant.
- **Experimental conditions:** Experimental conditions (including media, climatic conditions, disinfectant dosing procedure) are necessarily controlled and specific. How representative are these of true environmental conditions?
- **Definition of microbial inactivation:** Inactivation studies must select a criterion to define when a microorganism is damaged beyond recovery to an infectious state. Typically, loss of infectivity or culturability is the criterion for quantifying microbial inactivation. Possible additional criteria are usually based on the method to detect cell viability and include changes in electrical conductivity of cells, changes in the permeability of the cell membranes, presence of cell or enzyme activity, changes in the rate of metabolism and changes in the ability of the cell to stain with various dyes. For viruses, this can be integrity of the capsid and/or the nucleic acid. For protozoa, the ability to excyst (viability) has also been used. The most frequently used criterion for culturable organisms (most bacteria, several viruses, some protozoa) is complete loss of reproductive power. Some studies have explored the use of molecular methods to identify pathogen inactivation (Li et al., 2002; Shin & Sobsey, 2008; Lim, Kim & Ko, 2010). This approach relies on the assumption that genetic material is destroyed on pathogen inactivation and cannot be replicated during PCR. In reality, pathogens may be inactivated before genetic material is destroyed, and fragments of genes may be amplified during PCR. Therefore, molecular methods still typically overestimate infectious organism concentration and therefore underestimate microbial inactivation in comparison with cultural methods. However, active or viable but non-culturable (VBNC) states exist, meaning that pathogens may be infectious, yet not culturable (Lothigius et al., 2010). This VBNC state is also true for FIO (Sobsey et al., 1998; Juhna et al., 2007). Pathogen death or inactivation can also be investigated using infectivity studies (e.g. hepatitis A virus and murine norovirus) (Wei et al., 2010).

C7 References

- Ashbolt NJ, Grabow WOK, Snozzi M (2001). Indicators of microbial water quality. In: Fewtrell L, Bartram J, editors. Water quality: guidelines, standards and health – assessment of risk and risk management for water-related infectious disease. London: IWA Publishing on behalf of the World Health Organization; 289–315.
- Åström J, Petterson S, Bergstedt O, Pettersson TJ, Stenström TA (2007). Evaluation of the microbial risk reduction due to selective closure of the raw water intake before drinking water treatment. J Water Health. 5(Suppl 1):81–97.
- Badgley BD, Thomas FI, Harwood VJ (2010). The effects of submerged aquatic vegetation on the persistence of environmental populations of *Enterococcus* spp. Appl Environ Microbiol. 12:1271–81.
- Bastos RKX, Bevilacqua PD, Silva CAB, Silva CV (2008). Wastewater irrigation of salad crops: further evidence for the evaluation of the WHO guidelines. Water Sci Technol. 57(8):1213–9.
- Brookes JD, Antenucci J, Hipsey M, Burch MD, Ashbolt NJ, Ferguson C (2004). Fate and transport of pathogens in lakes and reservoirs. Environ Int. 30(5):741–59.
- Carlander A, Schönning C, Stenström TA (2009). Energy forest irrigated with wastewater: a comparative microbial risk assessment. J Water Health. 7(3):413–33.

- Craig DL, Fallowfield HJ, Cromar NJ (2003). Effectiveness of guideline faecal indicator organism values in estimation of exposure risk at recreational coastal sites. Water Sci Technol. 47(3):191–8.
- Davies CM, Mitchell VG, Petterson SR, Taylor GD, Lewis J, Kaucner C et al. (2008). Microbial challenge-testing of treatment processes for quantifying stormwater recycling risks and management. Water Sci Technol. 57(6):843–7.
- Diallo MB, Anceno AJ, Tawatsupa B, Houpt ER, Wangsuphachart V, Shipin OV (2008). Infection risk assessment of diarrhea-related pathogens in a tropical canal network. Sci Total Environ. 407(1):223–32.
- Diston D, Ebdon JE, Taylor HD (2012). The effect of UV-C radiation (254 nm) on candidate microbial source tracking phages infecting a human specific strain of *Bacteroides fragilis* (GB-124). J Water Health. 10:262–70.
- Dowd SE, Pillai SD (1997). A rapid viability assay for *Cryptosporidium* oocysts and *Giardia* cysts for use in conjunction with indirect fluorescent antibody detection. Can J Microbiol. 43:658–62.
- Eichmiller JJ, Hicks RE, Sadowsky MJ (2013). Distribution of genetic markers of fecal pollution on a freshwater sandy shoreline in proximity to wastewater effluent. Environ Sci Technol. 47:3395–402.
- Gilks WR, Richardson S, Spiegelhalter DJ (1996). Markov chain Monte Carlo methods in practice. London: Chapman and Hall.
- Gourmelon M, Caprais MP, Mieszkin S, Marti R, Wery N, Jarde E et al. (2010). Development of microbial and chemical MST tools to identify the origin of the faecal pollution in bathing and shellfish harvesting waters in France. Water Res. 44(16):4812–24.
- Grabow WOK (2001). Bacteriophages: update on application as models for viruses in water. Water SA. 27(2):251–68.
- Griffith JF, Cao Y, McGee CD, Weisberg SB (2009). Evaluation of rapid methods and novel indicators for assessing microbiological beach water quality. Water Res. 43:4900–7.
- Grimason AM, Smith HV, Thitai WN, Smith PG, Jackson MH, Girdwood RWA (1993). Occurrence and removal of *Cryptosporidium* spp. oocysts and *Giardia* spp. cysts in Kenyan waste stabilization ponds. Water Sci Technol. 27(3–4):97–104.
- Haas CN, Rose JB, Gerba CP (1999). Quantitative microbial risk assessment. Toronto: John Wiley & Sons.
- Hamilton AJ, Stagnitti F, Premier R, Boland AM, Hale G (2006). Quantitative microbial risk assessment models for consumption of raw vegetables irrigated with reclaimed water. Appl Environ Microbiol. 72(5):3284–90.
- Haramoto E, Otagiri M, Morita H, Kitajima M (2012). Genogroup distribution of F-specific coliphages in wastewater and river water in the Kofu basin in Japan. Lett Appl Microbiol. 54:367–73.
- Hata A, Kitajima M, Katayama H (2013). Occurrence and reduction of human viruses, F-specific RNA coliphage genogroups and microbial indicators at a full-scale wastewater treatment plant in Japan. J Appl Microbiol. 114:545–54.
- Haugland RA, Varma M, Sivaganesan M, Kelty C, Peed L, Shanks OC (2010). Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected Bacteroidales species and human fecal waste by qPCR. Syst Appl Microbiol. 33:348–57.
- Havelaar AH, De Hollander AEM, Teunis PFM, Evers EG, Van Kranen HJ, Versteegh JFM et al. (2000). Balancing the risks and benefits of drinking water disinfection: disability adjusted life-years on the scale. Environ Health Perspect. 108(4):315–21.
- Hendricks DW, Clunie WF, Sturbaum GD, Klein DA, Champlin TL, Kugrens P et al. (2005). Filtration removals of microorganisms and particles. J Environ Eng. 131(12):1621–32.
- Hijnen WAM, Dullemont YJ, Schijven JF, Hanzens-Brouwera AJ, Rosielled M, Medema G (2007). Removal and fate of *Cryptosporidium parvum*, *Clostridium perfringens* and small-sized centric diatoms (*Stephanodiscus hantzschii*) in slow sand filters. Water Res. 41:2151–62.
- Höglund C, Stenström TA, Ashbolt N (2002). Microbial risk assessment of source-separated urine used in agriculture. Waste Manage Res. 20(2):150–61.
- Howard G, Pedley S, Tibatemwa S (2006). Quantitative microbial risk assessment to estimate health risks attributable to water supply: can the technique be applied in developing countries with limited data? J Water Health. 4(1):49–65.
- Jaidi K, Barbeau B, Carriere A, Desjardins R, Prevost M (2009). Including operational data in QMRA model: development and impact of model inputs. J Water Health. 7(1):77–95.
- Jeanneau L, Solecki O, Wéry N, Jardé E, Gourmelon M, Communal P-Y et al. (2012). Relative decay of fecal indicator bacteria and human-associated markers: a microcosm study simulating wastewater input into seawater and freshwater. Environ Sci Technol. 46:2375–82.
- Juhna T, Birzniece D, Larsson S, Zulenkovs D, Sharipo A, Azevedo NF et al. (2007). Detection of *Escherichia coli* in biofilms from pipe samples and coupons in drinking water distribution networks. Appl Environ Microbiol. 73:7456–64.
- Ko G, Cromeans TL, Sobsey MD (2003). Detection of infectious adenovirus in cell culture by mRNA reverse transcription-PCR. Appl Environ Microbiol. 69(12):7377–84.

- LeChevallier M, Au K-K (2004). Water treatment and pathogen control: process efficiency in achieving safe drinking water. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/water-treatment-and-pathogen-control/en/, accessed 10 April 2016).
- LeChevallier MW, Norton WD, Lee RG (1991). *Giardia* and *Cryptosporidium* spp. in filtered drinking water supplies. Appl Environ Microbiol. 57(9):2617–21.
- Li JW, Xin ZT, Wang XW, Zheng JL, Chao FH (2002). Mechanisms of inactivation of hepatitis A virus by chlorine. Appl Environ Microbiol. 68(10):4951–5.
- Liang Z, He Z, Zhou X, Powell CA, Yang Y, Roberts MG et al. (2012). High diversity and differential persistence of fecal Bacteroidales population spiked into freshwater microcosm. Water Res. 46:247–57.
- Lim MY, Kim J-M, Ko G (2010). Disinfection kinetics of murine norovirus using chlorine and chlorine dioxide. Water Res. 44:3242–51.
- Lothigius Å, Sjöling Å, Svennerholm AM, Bölin I (2010). Survival and gene expression of enterotoxigenic *Escherichia coli* during long-term incubation in sea water and freshwater. J Appl Microbiol. 108:1441–9.
- Machdar E, van der Steen NP, Raschid-Sally L, Lens PNL (2013). Application of quantitative microbial risk assessment to analyze the public health risk from poor drinking water quality in a low income area in Accra, Ghana. Sci Total Environ. 449:134–42.
- Maier R, Pepper I, Gerba C (2009). Environmental microbiology, second edition. Academic Press.
- Makri A, Modarres R, Parkin R (2004). Cryptosporidiosis susceptibility and risk: a case study. Risk Anal. 24(1):209–20.
- Mara DD, Sleigh PA, Blumenthal UJ, Carr RM (2007). Health risks in wastewater irrigation: comparing estimates from quantitative microbial risk analyses and epidemiological studies. J Water Health. 5(1):39–50.
- Mara D, Hamilton A, Sleigh A, Karavarsamis N, Seidu R (2010). Tools for risk analysis: updating the 2006 WHO guidelines. In: Drechsel P, Scott C, Raschid-Sally L, Redwood M, Bahri A, editors. Wastewater irrigation and health. Assessing and mitigating risk in low-income countries. Colombo: International Water Management Institute; London: Earthscan; and Ottawa: International Development Research Centre.
- Masago Y, Katayama H, Hashimoto A, Hirata T, Ohgaki S (2002). Assessment of risk of infection due to *Cryptosporidium parvum* in drinking water. Water Sci Technol. 46(11–12):319–24.
- Masago Y, Oguma K, Katayama H, Hirata T, Ohgaki S (2004). *Cryptosporidium* monitoring system at a water treatment plant, based on waterborne risk assessment. Water Sci Technol. 50(1):293–9.
- Masago Y, Katayama H, Watanabe T, Haramoto E, Hashimoto A, Omura T et al. (2006). Quantitative risk assessment of noroviruses in drinking water based on qualitative data in Japan. Environ Sci Technol. 40(23):7428–33.
- McLellan SL, Newton RJ, Vandewalle JL, Shanks OC, Huse SM, Eren AM et al. (2013). Sewage reflects the distribution of human faecal Lachnospiraceae. Environ Microbiol. 15(8):2213–27.
- McQuaig S, Griffith J, Harwood VJ (2012). Association of fecal indicator bacteria with human viruses and microbial source tracking markers at coastal beaches impacted by nonpoint source pollution. Appl Environ Microbiol. 78:6423–32.
- McQuaig SM, Scott TM, Lukasik JO, Paul JH, Harwood VJ (2009). Quantification of human polyomaviruses JC virus and BK virus by TaqMan quantitative PCR and comparison to other water quality indicators in water and fecal samples. Appl Environ Microbiol. 75:3379–88.
- Medema GJ, Hoogenboezem W, van der Veer AJ, Ketelaars HAM, Hijnen WAM, Nobel PJ (2003). Quantitative risk assessment of *Cryptosporidium* in surface water treatment systems. Water Sci Technol. 47(3):241–7.
- Mitchell VG, O'Connor JL, Petterson S, Davies C, Kauncner C, Roser D et al. (2008). Quantifying stormwater recycling risks and benefits: data analysis and performance assessment. Final report to the Victorian Smart Water Fund.
- Oesterholt F, Martijnse G, Medema G, van der Kooij D (2007). Health risk assessment of non-potable domestic water supplies in the Netherlands. J Water Supply Res Technol – Aqua. 56(3):171–9.
- Oragui JI, Curtis TP, Silva SA, Mara DD (1986). The removal of excreted bacteria and viruses in deep waste stabilization ponds in northeast Brazil. Water Sci Technol. 18(10):31–5.
- Ottoson J, Stenström TA (2003). Faecal contamination of greywater and associated microbial risks. Water Res. 37(3):645–55.
- Payment P, Franco E (1993). *Clostridium perfringens* and somatic coliphages as indicators of the efficiency of drinking water treatment for viruses and protozoan cysts. Appl Environ Microbiol. 59(8):2418–24.
- Payment P, Locas A (2011). Pathogens in water: value and limits of correlation with microbial indicators. Ground Water. 49:4–11.
- Petterson SR, Stenström TA (2015). Quantification of pathogen inactivation efficacy by free chlorine disinfection of drinking water for QMRA. J Water Health. 13(3):625–44.

- Petterson SR, Signor RS, Ashbolt NJ (2007). Incorporating method recovery uncertainties in stochastic estimates of raw water protozoan concentration for QMRA. J Water Health. 5(Suppl 1):51–65.
- Petterson SR, Stenström TA, Ottoson J (2016). A theoretical approach to using faecal indicator data to model norovirus concentration in surface water for QMRA: Glomma River, Norway. Water Res. 91:31–7.
- Petterson SR, Teunis PFM, Ashbolt N (2001). Modeling virus inactivation on salad crops using microbial count data. Risk Anal. 21(6):1097–1107.
- Petterson SR, Dumoutier N, Loret JF, Ashbolt NJ (2009). Quantitative Bayesian predictions of source water concentration for QMRA from presence/absence data for *E. coli* O157:H7. Water Sci Technol. 59(11):2245–52.
- Petterson S, Grøndahl-Rosado R, Nilsen V, Myrmel M, Robertson LJ (2015). Variability in the recovery of a virus concentration procedure in water: implications for QMRA. Water Res. 87:79–86.
- Petterson SR, Mitchell VG, Davies CM, O'Connor J, Kaucner C, Roser D et al. (2016). Evaluation of three full-scale stormwater treatment systems with respect to water yield, pathogen removal efficacy and human health risk from faecal pathogens. Sci Total Environ. 543(Pt A):691–702.
- Polo C, Perez JL, Mielnichuck A, Fedele CG, Niubo J, Tenorio A (2004). Prevalence and patterns of polyomavirus urinary excretion in immunocompetent adults and children. Clin Microbiol Infect. 10:640–4.
- Pouillot R, Beaudeau P, Denis JB, Derouin F (2004). A quantitative risk assessment of waterborne cryptosporidiosis in France using second-order Monte Carlo simulation. Risk Anal. 24(1):1–17.
- Regli S, Rose JB, Haas CN, Gerba CP (1991). Modeling the risk from *Giardia* and viruses in drinking water. J Am Water Works Assoc. 83(11):76–84.
- Rose JB, Dickson LJ, Farrah SR, Carnahan RP (1996). Removal of pathogenic and indicator microorganisms by a full-scale water reclamation facility. Water Res. 30(11):2785–97.
- Roser D, Ashbolt N, Storey M, Leeming R, Scott P, Bennetts R et al. (2002). Application of physico-chemical parameters and particle-bound biomarkers to indicate microbial contamination of aquifers. Water Sci Technol Water Supply. 2(5–6):419–26.
- Schijven JF, de Roda Husman AM (2011). Applications of quantitative microbial source tracking (QMST) and quantitative microbial risk assessment (QMRA). In: Hagedorn C, Blanch A, Harwood J, editors. Microbial source tracking: methods, applications, and case studies. Springer Science+Business Media, LLC; 559–83.
- Schijven JF, Hassanizadeh SM (2000). Removal of viruses by soil passage: overview of modeling, processes, and parameters. Crit Rev Environ Sci Technol. 30(1):49–127.
- Schijven JF, Hoogenboezem W, Hassanizadeh SM, Peters JH (1999). Modeling removal of bacteriophages MS2 and PRD1 by dune recharge at Castricum, Netherlands. Water Resour Res. 35(4):1101–11.
- Schijven JF, de Bruin HAM, Hassanizadeh SM, de Roda Husman AM (2003). Bacteriophages and *Clostridium* spores as indicator organisms for removal of pathogens by passage through saturated dune sand. Water Res. 37(9):2186–94.
- Schijven JF, Mulschlegel JH, Hassanizadeh SM, Teunis PF, de Roda Husman AM (2006). Determination of protection zones for Dutch groundwater wells against virus contamination – uncertainty and sensitivity analysis. J Water Health. 4:297–312.
- Schoen ME, Soller JA, Ashbolt NJ (2011). Evaluating the importance of faecal sources in human-impacted waters. Water Res. 45:2670–80.
- Schwartzbrod L (1995). Effect of human viruses on public health associated with the use of wastewater and sewage sludge in agriculture and aquaculture. Geneva: World Health Organization (WHO/EOS/95.19; http://apps.who.int/iris/handle/10665/59685, accessed 5 May 2016).
- Scott TM, McLaughlin MR, Harwood VJ, Chivukula V, Levine A, Gennaccaro A et al. (2003). Reduction of pathogens, indicator bacteria, and alternative indicators by wastewater treatment and reclamation processes. Water Sci Technol Water Supply. 3(4):247–52.
- Seidu R, Heistad A, Amoah P, Drechsel P, Jenssen PD, Stenström TA (2008). Quantification of the health risk associated with wastewater reuse in Accra, Ghana: a contribution toward local guidelines. J Water Health. 6(4):461–71.
- Shanks OC, Kelty CA, Sivaganesan M, Varma M, Haugland RA (2009). Quantitative PCR for genetic markers of human fecal pollution. Appl Environ Microbiol. 75:5507–13.
- Shin G-A, Sobsey MD (2008). Inactivation of norovirus by chlorine disinfection of water. Water Res. 42:4562–8.
- Shuval H, Lampert Y, Fattal B (1997). Development of a risk assessment approach for evaluating wastewater reuse standards for agriculture. Water Sci Technol. 35(11–12):15–20.
- Signor RS, Ashbolt NJ (2006). Pathogen monitoring offers questionable protection against drinking-water risks: a QMRA (quantitative microbial risk analysis) approach to assess management strategies. Water Sci Technol. 54(3):261–8.

- Signor RS, Roser DJ, Ashbolt NJ, Ball JE (2005). Quantifying the impact of runoff events on microbiological contaminant concentrations entering surface drinking source waters. J Water Health. 3(4):453–68.
- Smeets PW (2008). Stochastic modeling of drinking water treatment in quantitative microbial risk assessment. Delft: Water Management Academic Press.
- Smeets PW, van Dijk JC, Stanfield G, Rietveld LC, Medema GJ (2007). How can the UK statutory *Cryptosporidium* monitoring be used for quantitative risk assessment of *Cryptosporidium* in drinking water? J Water Health. 5(Suppl 1):107–18.
- Sobsey MD (1989). Inactivation of health-related microorganisms in water by disinfection processes. Water Sci Technol. 21(3):179–95.
- Sobsey MD, Battigelli DA, Shin G-A, Newland S (1998). RT-PCR amplification detects inactivated viruses in water and wastewater. Water Sci Technol. 38(12):91–4.
- Soller JA, Eisenberg JNS, DeGeorge JF, Cooper RC, Tchobanoglous G, Olivieri AW (2006). A public health evaluation of recreational water impairment. J Water Health. 4(1):1–19.
- Staley C, Gordon KV, Schoen ME, Harwood VJ (2012). Performance of two human-associated microbial source tracking qPCR methods in various Florida water types and implications for microbial risk assessments. Appl Environ Microbiol. 78:7317–26.
- Teunis PFM, Evers EG, Slob W (1999). Analysis of variable fractions resulting from microbial counts. Quant Microbiol. 1:63–88.
- Teunis PFM, Ogden ID, Strachan NJC (2008). Hierarchical dose response of *E. coli* O157:H7 from human outbreaks incorporating heterogeneity in exposure. Epidemiol Infect. 136:761–70.
- Teunis PFM, Medema GJ, Kruidenier L, Havelaar AH (1997). Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. Water Res. 31(6):1333–46.
- Teunis PFM, Rutjes SA, Westrell T, de Roda Husman AM (2009). Characterization of drinking water treatment for virus risk assessment. Water Res. 43(2):395–404.
- Toledo-Hernandez C, Ryu H, Gonzalez-Nieves J, Huertas E, Toranzos GA, Santo Domingo J (2013). Tracking the primary sources of fecal pollution in a tropical watershed: a one year study. Appl Environ Microbiol. 79:1689–96.
- van Heerden J, Ehlers MM, Vivier JC, Grabow WO (2005). Risk assessment of adenoviruses detected in treated drinking water and recreational water. J Appl Microbiol. 99(4):926–33.
- van Lieverloo JH, Blokker EJ, Medema G (2007). Quantitative microbial risk assessment of distributed drinking water using faecal indicator incidence and concentrations. J Water Health. 5(Suppl 1):131–49.
- Vose D (2008). Risk analysis: a quantitative guide. John Wiley & Sons.
- Wade TJ, Sams E, Brenner KP, Haugland R, Chern E, Beach M et al. (2010). Rapidly measured indicators of recreational water quality and swimming-associated illness at marine beaches: a prospective cohort study. Environ Health. 9:66. doi:10.1186/1476-069X-9-66.
- Walpole RE, Myers RH, Myers SL, Ye K (2011). Probability and statistics for engineers and scientists, ninth edition. Prentice Hall.
- Wei J, Jin Y, Sims T, Kniel KE (2010). Survival of murine norovirus and hepatitis A virus in different types of manure and biosolids. Foodborne Pathog Dis. 7:901–6.
- Westrell T, Bergstedt O, Stenström TA, Ashbolt NJ (2003). A theoretical approach to assess microbial risks due to failures in drinking water treatment. Int J Environ Health Res. 13(2):181–97.
- Westrell T, Teunis P, van den Berg H, Lodder W, Ketelaars H, Stenström TA et al. (2006). Short and long term fluctuations of norovirus concentrations in surface water and their implications for public health. Water Res. 40(14):2613–20.
- WHO (1996). Guidelines for drinking-water quality. Vol. 2. Health and other supporting criteria, second edition. Geneva: World Health Organization.
- WHO (2006). Guidelines for the safe use of wastewater, excreta and greywater. Vol. 2. Wastewater use in agriculture. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/gsuweg2/en/, accessed 9 April 2016).
- WHO (2011). Evaluating household water treatment options: health-based targets and microbiological performance specifications. Geneva: World Health Organization (http://www.who.int/water_sanitation_health/publications/household_water/en/, accessed 10 April 2016).
- WHO (in preparation). Guidelines for drinking-water quality, fourth edition incorporating first addendum. Geneva: World Health Organization.

ANNEX D | DOSE–RESPONSE

Most risk assessments use dose–response information from the published literature. Care must be taken that the selected model and parameter values are appropriate for the risk study and that when interpreting the calculations, the implications of the assumptions underlying the selected numbers are considered.

The infectivity of a pathogen depends on its ability to pass host defences, find a site for colonization and cause an infection within the human host. Variability in infectivity can then exist due to:

- variability between pathogens in their virulence and human infectivity; and
- variability between hosts in the strength of their immune response. Susceptible portions of the population include children, the elderly and the immunocompromised, who are more easily infected. In addition, short- or long-term immunity following an infection may protect portions of the population from future enteric infections.

D1 Observations

Two types of studies are used to obtain scientific evidence to support the development of dose–response relationships: clinical studies and outbreaks. In all cases, the following factors need to be considered in interpreting the data:

- How many subjects (observations) were included in the study?
- Who were the study subjects, and what was their vulnerability to disease (i.e. animals, healthy children and/or healthy adults, or immunocompromised or malnourished children and/or adults)?
- What was the history of the organism used for the experiment?
- What were the exposure doses?
- How was infection determined?

D1.1 Clinical studies

In clinical studies, subjects are administered a known dose of a laboratory strain of pathogens, and their response is then followed over days or weeks following exposure for evidence of infection (excretion, serological response) and illness. The majority of dose–response data used for QMRA are from healthy adult volunteers, although animal studies are also used in some specific cases.

D1.2 Outbreaks

In some cases, real outbreaks have been used to quantify the dose–response relationship for certain pathogens. The advantages of outbreak studies are that the pathogen is a true environmental strain and the population is a

true representative population (often including children). The disadvantages are that the dose is often poorly known and the response of all exposed subjects is difficult to follow.

D2 Dose–response models

Quantifying infectivity involves selecting a model to describe the process and fitting that model to the observations in order to obtain best estimates of the parameter values.

Various models exist for the dose–response relationship for infection and the dose–response relationship for illness when infected (the conditional relation for illness given infection). Most are based on the single-hit theory: where every ingested pathogen particle is assumed to act independently and has an individual probability of causing infection (Haas, 1983; Teunis & Havelaar, 2000). Population dose–response models for infection are summarized in Box D.1. Note that there are two different forms of the Beta-Poisson model, an exact and an approximate function. Care should be taken that when parameter estimates from the exact model are not valid for the approximate model ($\alpha < \beta$ and $\beta > 1$), these parameter estimates should *not* be used for predicting risks with the approximate model, because they will result in serious errors in the calculated risk.

Box D.1 Dose–response relationship for infection

For a Poisson inoculum, the dose-dependent probability of exposure is:

$$P_{inf}\,(c{\cdot}V) = 1 - e^{-c \cdot V} \qquad \text{Eq. D.1}$$

with dose cV as the concentration of pathogens (c) multiplied by ingested volume (V). If the single-hit probability of infection (probability that any single ingested pathogen succeeds in infecting the host) is r, then the dose–response model for infection is:

$$P_{inf}\,(c{\cdot}V\,;\,r) = 1 - e^{-rc \cdot V} \qquad \text{Eq. D.2}$$

If r is variable and its variability is described by a beta distribution with parameters (α, β), then:

$$P_{inf}\,(c{\cdot}V\,;\,\alpha,\beta) = 1 - {}_1F_1(\alpha, \alpha+\beta; -c{\cdot}V) \qquad \text{Eq. D.3}$$

in which ${}_1F_1$ is a confluent hypergeometric function (Kummer function). Provided that $\alpha \ll \beta$ and $\beta \gg 1$, this model, called the "Beta-Poisson" dose–response model, can be approximated by:

$$P_{inf}\,(c{\cdot}V\,;\,\alpha,\beta) = 1 - \left(1 + \frac{c{\cdot}V}{\beta}\right)^{-\alpha} \qquad \text{Eq. D.4}$$

This latter approximation is sometimes written as

$$P_{inf}\,(c{\cdot}V\,;\,\alpha,N_{50}) = 1 - \left(1 + \frac{c{\cdot}V}{N_{50}}\left(2^{1/\alpha} - 1\right)\right)^{-\alpha} \qquad \text{Eq. D.5}$$

in which N_{50} is the ID_{50}, the 50% infectious dose (Haas et al., 1993). Note that N here is not a discrete number: the dose is a Poisson parameter and can be any positive real number.

Several modifications to the dose–response models have been published, allowing for covariables such as immune status of the host, innate immunity, aggregated pathogens or sexual reproduction of the pathogen. This (conditional) probability may also depend on dose (Teunis, Nagelkerke & Haas, 1999; Teunis et al., 2005, 2010). This has consequences for epidemiology, because when a population is exposed to low doses of an enteric pathogen, some people may become infected, but few of these infected subjects will become ill, so that cases are sporadic. When the dose is high, not only is there a higher probability of infection, but also those who are infected have a higher probability of becoming ill, so that a cluster of enteric illness may be detected.

D2.1 Maximum risk curve: the upper limit of the single-hit model

A useful property of the single-hit model is that there is an upper bound, referred to as the maximum risk curve (Teunis & Havelaar, 2000). This curve describes the probability of infection when every ingested organism is assumed to be successful in causing an infection (exponential model with $r = 1$). In the absence of pathogen-specific infectivity information or to estimate the worst case for a particularly virulent pathogen strain, the maximum risk curve can be used as an upper bound.

D2.2 Fractional Poisson model

In the fractional Poisson model proposed by Messner, Berger & Nappier (2014), the individual's probability of infection (r) is assumed to be exactly 1 with probability P and exactly 0 with probability $1 - P$, and it cannot be anything in between. The probability of infection (P_{inf}) is given by:

$$P_{inf} = P \cdot (1 - e^{-\frac{dose}{\mu}}) \qquad \text{Eq. D.6}$$

where μ is the mean aggregated size (as the model was presented in the context of norovirus). Individuals are therefore considered to be perfectly susceptible or perfectly immune. The probability of receiving at least one aggregate (and hence at least one virion) is given by $(1 - e^{-\frac{dose}{\mu}})$, which may also be referred to as the probability of exposure.

D2.3 Application of single-hit models in QMRA: adopting a conditional approach

When applying single-hit models, the distribution of pathogens in exposure media is assumed to be random (Poisson). For human volunteer feeding trial studies, the Poisson distribution may be a reasonable assumption; however, in most environmental media, some microbial clumping (overdispersion) is expected. When applying dose–response models for QMRA and using Monte Carlo simulation, rather than using the exponential or Beta-Poisson models (with input as average dose), it may be preferred to simulate the number of organisms consumed by each exposed individual within the Monte Carlo and then calculate the probability of infection based on that dose (Haas, 2002). The probability that each organism is capable of causing infection is given by r (for the exponential model) or a beta distribution defined by α and β (for Beta-Poisson or hypergeometric models). Using this approach, the pathogens can be assumed to follow any distribution in environmental media.

D2.4 Application of single-hit models in QMRA: low-dose approximations

In the implementation of dose–response models within QMRA, it is often desirable to apply a low-dose approximation in order to simplify the calculations. In the low-dose region, the relationships are close to linear, and, in many cases, an approximation is reasonable. In the application of dose–response models in the drinking-water guidelines (GDWQ Table 7.4; WHO, in preparation), a low-dose simplification to the single-hit model has been adopted. Only the probability of infection from a single organism is considered; the probability of exposure to more than one organism per day is assumed to be so low that it is ignored in the dose–response relationship. At very low concentrations, this simplification is valid; however, it is important that at higher pathogen doses (approximately 0.3 microorganisms), the full single-hit (exponential or Beta-Poisson) model should be used to account for the risk of infection associated with exposure to more than one organism.

When applying the Beta-Poisson model, the commonly used approximation ($P_{inf} = 1 - (1 + \frac{dose}{\beta})^{-\alpha}$) (where α and β are parameters of the beta distribution) is appropriate when the parameter values meet the constraints of the approximation ($\alpha << \beta$ and $\beta >> 1$). As illustrated in Fig. D.1, for parameter values that do not lie in this range, the common approximation can deviate considerably from the exact Beta-Poisson model and may even (as in the case of the *Campylobacter jejuni* model) exceed the theoretical maximum risk. In these cases, a better low-dose approximation is to use the exponential dose–response function with r equal to the expected value of the beta distribution ($\frac{\alpha}{\alpha + \beta}$).

D3 Published dose–response relationships

Many studies have been undertaken in which dose–response models have been fitted to experimental data. Published parameter values from these studies have been widely applied within QMRA in a diverse range of contexts. Each of these models has different sources of uncertainty, as they are based on different data (different doses, different measures of dose, different hosts and different sample sizes), and modelling studies have applied a range of different assumptions to produce an acceptable fit. An appreciation of these details is valuable for selecting a dose–response model and subsequently interpreting the representativeness of the resulting risk calculations.

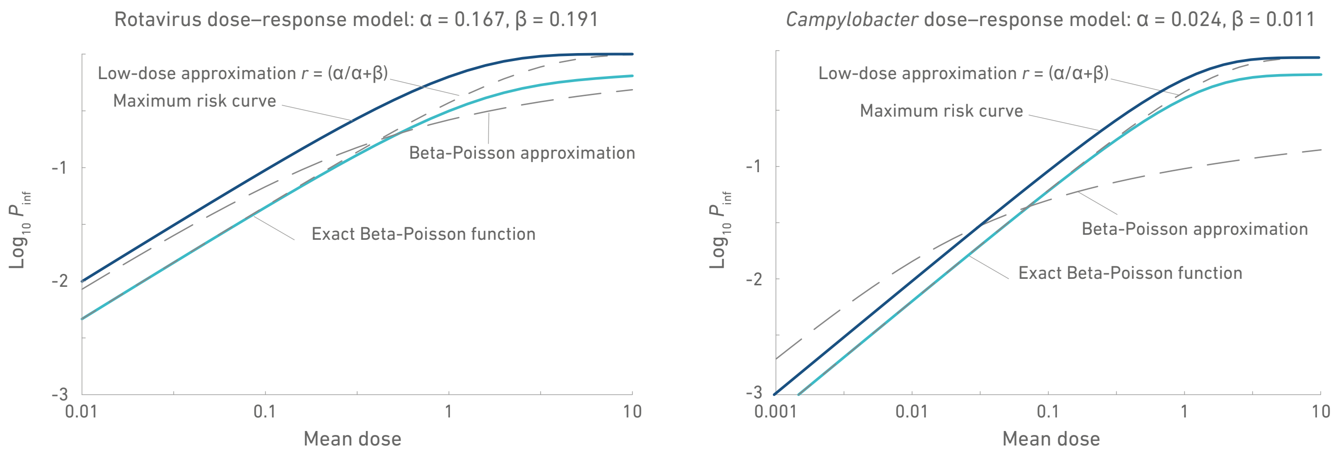


Fig. D.1 Comparison of low-dose approximations for rotavirus and *Campylobacter* models (adapted from Medema et al., 2006)

D3.1 *Campylobacter jejuni*

Medema and colleagues (1996) developed a dose–response relationship for *Campylobacter jejuni* by fitting the Beta-Poisson approximation to experimental data published by Black and coworkers (1988). In that study, healthy adult volunteers were fed doses of between 8 × 10^2 and 1 × 10^8 cells prepared with milk. In total, 68 volunteers were included, with between five and 19 volunteers challenged with each of six doses. Infection was defined as a positive result from stool culture; a subset of the infected volunteers also developed symptoms of illness.

The dosage and subsequent attack rates associated with the volunteer study were very high, with the lowest experimental dose at 800 CFU. The subsequent dose–response relationship has been extensively applied for water-related QMRAs where the mean doses are often very low, much lower than the experimental range. Indeed, when the objective of the QMRA is to model the annual risk of infection in the range of 1 × 10^{-4} (with daily exposure), then the mean daily dose must be approximately 1 × 10^{-5} CFU, around 7 $\log_{10}$ below the range of the experimental observations on which the dose–response relationship was based.

In a later study, Teunis and coworkers (2005) combined the experimental data from the volunteer study with observations of illness from two milk-related outbreaks: one in the United Kingdom and another in the Netherlands. In both outbreaks, a clear relationship was observed between the attack rate and the amount of milk consumed; however, the concentration of bacteria in the contaminated milk was not known. A statistical modelling approach was implemented to estimate the most likely parameter values of the Beta-Poisson distribution, given both the outbreaks and the volunteer study. The concentration of *C. jejuni* in the contaminated milk was assumed to be the same for both outbreaks and was included as a parameter in the model. For a given concentration, all other parameters in the model could be determined. The authors noted a local peak in likelihood in the region of about 15 CFU·L^{-1}, and therefore α and β values were determined at this concentration. The information from the outbreak “filled in” the gap in the dose–response relationship from the human volunteer study, providing information at lower doses, although still much higher than used in application.

The resulting updated model predicts a much higher infectivity at low doses than the previous model (expected value of r = 0.69 in comparison with r = 0.019), which has important implications for practical implementation. For example, using the original dose–response relationship, the health-based treatment target for *Campylobacter* in the WHO GDWQ is 5.9 $\log_{10}$ removal (WHO, in preparation). However, if the updated model were to be implemented, the treatment target would increase to approximately 7.5 $\log_{10}$. Would this increase be justified, or is it overly conservative?

Since publication in 2005, there has been mixed uptake of the updated model, with some authors adopting it (Åström et al., 2007; van Lieverloo, Blokker & Medema, 2007; Soller et al., 2010) and others favouring the earlier dose–response relationship developed by Medema et al. (1996) (e.g. Mara et al., 2007; Oesterholt et al., 2007; Bastos et al., 2008; Smeets et al., 2008). The higher infectivity and the dependency of the predicted model parameters on the unknown concentration in contaminated milk are both identified as contributing causes to the lack of uptake. The uncertainty in defining dose is greater in the outbreak studies; however, the strain of *C. jejuni* is indeed a true environmental strain (as opposed to a laboratory strain), and the population exposed included

children. Both of these factors may be expected to contribute to a higher prediction of infectivity, and both are relevant for real populations. While the infectivity of the updated model may appear too high, it is based on a broader scientific evidence base than is reflected by the volunteer study alone.

D3.2 *E. coli* O157:H7

Clinical data on the infectivity of *E. coli* O157:H7 to humans are not available. Dose–response relationships have been developed from animal studies, surrogate organisms and outbreaks:

- **Animal studies:** Pai, Kelly & Meyers (1986) studied the infectivity of *E. coli* O157:H7 in infant New Zealand White rabbits. In total, 39 rabbits were fed doses of organisms ranging from 10^5 to 10^{10} cells, and 32 rabbits were infected. Haas and coworkers (2000) fitted the Beta-Poisson approximation to these data ($\alpha = 0.49$; $N_{50} = 5.9 \times 10^5$) and concluded that the predictions of the model were concordant with illness rates evidenced in two documented outbreaks.
- **Surrogate organisms:** Powell et al. (2000) combined human feeding study results from two surrogate organisms, enteropathogenic *E. coli* and *Shigella dysenteriae*. The authors argued that given the biological understanding of pathogenesis, the clinical trial data and the epidemiological information, it was reasonable to infer that *E. coli* O157:H7 has an effective dose somewhat higher than that of *S. dysenteriae* and somewhat lower than that of enteropathogenic *E. coli*. They fitted a model including clinical data for the two surrogates to define an uncertainty envelope for the dose–response relationship for *E. coli* O157:H7.
- **Outbreaks:** Teunis et al. (2004) analysed data from a single outbreak in which schoolchildren and teachers ate contaminated food (Shinagawa, Hu & Yoshida, 1997). Of the 828 children and 43 adults who were exposed, 208 and seven, respectively, were found infected. This later study compared their results with the surrogate study and animal study and argued that the results from *Shigella dysenteriae* (analysed by Teunis et al., 1996) appeared to have the greatest agreement with the actual outbreak data (Teunis, Takumi & Shinagawa, 2004). Teunis, Ogden & Strachan (2008) later extended this approach to examine eight different outbreaks; altogether, 14 654 individuals were exposed, and 718 became infected. A hierarchical model (a beta distribution fit to infectivity within each outbreak; and another beta distribution to describe the distribution between outbreaks) was applied to allow for between-outbreak variability separately and incorporated overdispersion (relative to the Poisson distribution) in dose, with the resulting beta distribution for infectivity defined by medians of the posterior distributions for $\alpha = 0.373$ and $\beta = 39.71$.

D3.3 Enteroviruses

D3.3.1 Echovirus-12

Schiff and coworkers (1984) investigated the infectivity of echovirus-12 on 149 healthy adults (18–45 years of age). The virus was isolated from a child with erythema infectiosum (also known as fifth disease or slapped cheek syndrome) and passed twice in primary rhesus monkey kidney cells before each volunteer was exposed to 0–330 000 PFU of virus. Infection was defined as faecal shedding or significant increases in serum antibody titre, and no subjects developed significant illness. Teunis et al. (1996) fitted both the exponential and Beta-Poisson approximation models to the clinical data set. The Beta-Poisson approximation provided the better fit ($\alpha = 0.401$; $\beta = 227.2$). These data have been used for QMRA of drinking-water to represent the infectivity of the enteroviruses (Regli et al., 1991; Åström et al., 2007; van Lieverloo, Blokker & Medema, 2007).

D3.3.2 Coxsackievirus B4

Suptel (1963) undertook dose–response experiments with mice using coxsackievirus B4. In total, 50 mice were exposed to doses of 1.3–13 000 PFU, with 31 becoming infected (mouse mortality was taken as the indication of infection). Mena and colleagues (2003) eliminated the data at the lowest dose of 1.3 PFU (in order to make a model fit possible) and fitted both the exponential and Beta-Poisson approximations to the data set. No significant improvement in fit was achieved by the Beta-Poisson model, and therefore the exponential model was selected ($r = 0.0078$). Mena and coworkers (2003) applied this model directly to characterize the risk associated with exposure of humans to coxsackievirus via drinking-water and recreational water.

D3.4 Adenovirus

Clinical data that have been used to model adenovirus infectivity for QMRA were obtained as part of an aerosol disease transmission study. Subjects were nine healthy adult male inmates who were exposed to doses of 1, 5 or 11 $TCID_{50}$ of adenovirus type 4 via aerosol particles (0.3–2.5 μm). Seven individuals became infected. The two individuals who resisted infection were exposed to the lowest dose (1 $TCID_{50}$), indicating an increasing dose–response relationship over the range. An exponential dose–response model was fitted to these data ($r = 0.4172$), and it has subsequently been used for risk assessment of enteric adenovirus via ingestion for a range of water-related enteric exposure scenarios (Crabtree et al., 1997; van Heerden et al., 2004; Westrell et al., 2004).

D3.5 Rotavirus

Clinical data for rotavirus infectivity are drawn from a human feeding study published by Ward, Bernstein & Young (1986), in which 62 adult men were administered doses of 9×10^{-3} to 9×10^{4} FFU of rotaviruses sourced from the faeces of a sick child (8 years old). In total, 30 subjects became infected (excretion of rotavirus or seroconversion, or both), and 17 showed signs of illness. These data were among the first to be used for QMRA with application of the Beta-Poisson approximation model ($\alpha = 0.26$; $\beta = 0.42$) (Regli et al., 1991; Haas et al., 1993).

Teunis & Havelaar (2000) identified an important limitation associated with the application of this model – that is, that the conditions of the Beta-Poisson approximation were not met, leading to a poor approximation at low doses. Indeed, during uncertainty analysis, Teunis & Havelaar (2000) demonstrated that the credible interval for the dose–response relationship exceeded the theoretical maximum risk. Teunis & Havelaar (2000) therefore fit the exact hypergeometric Beta-Poisson model to the clinical data ($\alpha = 0.167$; $\beta = 0.191$).

When applying either of these models for QMRA, it is relevant to consider that the doses for the clinical study were measured in FFU. It is therefore important to consider the units of exposure – if exposure is quantified in terms of molecular units (e.g. genome copies or PDU), then a conversion is required before using the dose–response relationship.

D3.6 Norovirus

Clinical data are available from three separate experiments that were undertaken with two different inocula (Teunis et al., 2008). Two experiments were undertaken with the first inoculum, which had been stored in a high-protein (sticky) suspension for more than 25 years and was highly aggregated. The third experiment was undertaken with the second inoculum, which was extracted from a stool sample of one of the infected subjects from the first experiment and contained in phosphate-buffered saline. In total, 80 healthy human volunteers (secretor positive and considered susceptible to infection) were challenged with doses between 3.24×10 and 3.24×10^{8} genomes, of which 40 became infected and 24 (±1) became ill.

Teunis et al. (2008) modelled these data to estimate parameter values for the dose–response relationship; however, a challenge was accounting for the differences in the inocula. The first inoculum was aggregated, and the second was not; and the second inoculum had passed through a human, which may have influenced its infectivity. As expected, when the dose–response models were fitted to each of the data sets separately, the resulting dose–response relationships were (statistically) different.

Teunis and coworkers (2008) hypothesized that the difference between the relationships may have been primarily due to aggregation state and therefore developed a modified single-hit model that allowed for shared α and β values (e.g. identical infectivity between data sets), but different aggregation states. Using a likelihood ratio test to show equivalent fit of the combined model in comparison with the separate models, the authors concluded that the difference between the two experiments could be explained by aggregation. The implication of this result is that the observations from all 80 volunteers were used to infer the parameter values of the beta distribution describing norovirus infectivity ($\alpha = 0.04$; $\beta = 0.055$), and the influence of aggregation on the results was eliminated. These parameter values can be applied in the exact hypergeometric Beta-Poisson model for QMRA.

In the same manner as for the rotavirus model, it is necessary to ask, what are the appropriate units of "dose" for this dose–response relationship? The doses for the clinical study were described as the number of genomes; however, the statistical modelling undertaken by Teunis et al. (2008) accounted for aggregation of genomes within virus particles. In order to translate a molecular result from an environmental sample (e.g. number of genome copies per litre of water) to an exposure dose for the dose–response relationship, it is necessary to know

something about the aggregation state of the sample. In practice, this is not dissimilar from the rotavirus model that relates probability of infection to the number of virus aggregates (FFU) that were consumed.

More recently, Messner, Berger & Nappier (2014) combined the data reported by Teunis et al. (2008) with three additional human challenge studies (Seitz et al., 2011; Frenck et al., 2012; Atmar et al., 2013). They fitted the exact Beta-Poisson model (accounting for aggregation) to the combined data set, providing updated parameter values (Table D.1). In addition, they proposed a new dose–response model for human norovirus, which they referred to as the fractional Poisson model. In the fractional model, each individual is assumed to have a probability of exactly 0 or 1 of becoming infected. Comparison with the exact Beta-Poisson model fitted by Teunis et al. (2008) using the AIC indicated that not only was the fractional Poisson model computationally simpler than the Beta-Poisson, but the model fit was superior.

D3.7 *Giardia*

Forty adult male prison inmates were administered one of eight doses containing between 1 and 10^6 *Giardia* cysts and monitored for infection. A positive response was determined as the presence of viable *Giardia lamblia* cysts in faeces. The exponential model was fitted to the data set using the method of maximum likelihood, with an MLE for $r = 0.0198$ (Rose et al., 1991), and it has been widely used in water-related QMRA (e.g. Regli et al., 1991; Ottoson & Stenström, 2003; Gale, 2005; Åström et al., 2007; Diallo et al., 2008).

D3.8 *Cryptosporidium*

A series of human feeding trials using healthy adult volunteers have been undertaken for *Cryptosporidium parvum*, *C. hominis*, *C. meleagridis* and *C. muris*. The first, using the *C. parvum* Iowa isolate propagated in calves, was published in 1995 (DuPont et al., 1995). Many risk assessments were subsequently undertaken using these data, to which the exponential model was fitted ($r = 0.0042$) (e.g. Haas et al., 1996; Teunis et al., 1997; Masago et al., 2002; Medema et al., 2003).

Subsequent trials were undertaken with the TAMU and UCP (Okhuysen et al., 1999) and Moredun isolates (Okhuysen et al., 2002), indicating considerable between-isolate variability in *Cryptosporidium* infectivity. Teunis, Chappell & Okhuysen (2002) fitted the exponential and Beta-Poisson single-hit models to the Iowa, TAMU and UCP data sets. The first two data sets were fitted adequately by the exponential model; only the UCP data set demonstrated a significant improvement in fit with the hypergeometric model. Comparison between the exponential parameter (r) for the Iowa and TAMU data sets (Table D.1) indicated around a 1 $\log_{10}$ higher probability of infection from a single organism for TAMU in comparison with Iowa, and around another $\log_{10}$ increase (based on the expected value of the beta distribution) for UCP in comparison with Iowa. This appears to be a highly significant difference in infectivity; however, it is important to note that Teunis, Chappell & Okhuysen (2002) identified only a slight statistically significant improvement in model fit by analysing the three data sets separately, rather than pooling them together as a combined data set. How is it possible that if the infectivity is so vastly different, it is not statistically essential to consider the data sets separately? It is possible because of the magnitude of uncertainty associated with analysing the clinical data, and hence with the predicted parameter values for dose–response. We cannot be certain that the infectivity is so different between isolates, but, given the clinical observations, it is likely. This example of *Cryptosporidium* clearly highlights the scale of the uncertainty that needs to be considered for all pathogen data sets; the data sets are so small, and there is so much uncertainty underlying the observations, that quantitative conclusions about the magnitude of infectivity are a challenge.

Table D.1 Summary of single-hit dose–response relationships commonly used for water-related QMRA

Reference pathogen	Observational data			
	N	Minimum dose	Maximum dose	Reference
Campylobacter	68	8×10^2 cells	1×10^8 cells	Black et al. (1988)
		Unknown		Evans et al. (1996); van den Brandhof, Wagenaar & van den Kerkhof (2003)
E. coli O157:H7				
Rabbits	39	1×10^5	1×10^{10}	Pai, Kelly & Meyers (1986)
Outbreaks	14 654	14	1.0×10^4	Various outbreaks
Enteroviruses				
Coxsackievirus B4 (mice)	50	1.3	13 000 PFU	Suptel (1963)
Echovirus-12	149	0	330 000 PFU	Schiff et al. (1984)
Norovirus	80	3.24 gec	3.24×10^8 gec	Teunis et al. (2008)
	157	3.24 gec	3.24×10^8 gec	Teunis et al. (2008); Seitz et al. (2011); Frenck et al. (2012); Atmar et al. (2013)
	157	3.24 gec	3.24×10^8 gec	Teunis et al. (2008); Seitz et al. (2011); Frenck et al. (2012); Atmar et al. (2013)
Rotavirus	62	9×10^{-3}	9×10^4 FFU	Ward, Bernstein & Young (1986)
Adenoviruses				
Adenovirus 4 exposure via inhalation	9	1	11 $TCID_{50}$	Couch et al. (1966)
Giardia	40	1	10^6	Rendtorff (1954)
Cryptosporidium parvum				
Iowa	29	30	10^6	DuPont et al. (1995)
Iowa	29	30	10^6	DuPont et al. (1995)
TAMU	14	10	500	Okhuysen et al. (1999)
UCP	17	500	10 000	Okhuysen et al. (1999)
Combined (Iowa, TAMU, UCP)	60	10	10^6	Okhuysen et al. (1999)
Cryptosporidium hominis	21	10	500	Chappel et al. (2006)

Approx. BP: Beta-Poisson approximation; Exact BP: exact Beta-Poisson model; Exp.: exponential; gec: genomic equivalent copies; N_{50}: an alternative parameterization of the Beta-Poisson model

[a] Expected value of $r = \alpha/(\alpha+\beta)$.

[b] Probability of being perfectly susceptible; see section D2.1.

[c] These studies analysed the same data set using different methods of parameter fitting.

Dose–response relationship					
Model	α	β	*r*	$E(r)$[a]	Reference
Approx. BP	0.145	7.59	–	0.019	Medema et al. (1996)
Exact BP	0.024	0.011	–	0.69	Teunis et al. (2005)
Approx. BP	0.49	N_{50}: 5.9×10^5	–	–	Haas et al. (2000)
Exact BP	0.373	39.71	–	0.009 3	Teunis, Ogden & Strachan (2008a)
Exp.	–	–	0.007 75	–	Mena et al. (2003)
Approx. BP	0.401	227.2	–	0.001 8	Teunis et al. (1996)
Exact BP	0.040	0.055	–	0.42	Teunis et al. (2008)
Exact BP	0.006 3	0.003 2	–	0.66	Messner, Berger & Nappier (2014)
Fractional Poisson	–	–	1 or 0 $P = 0.72$[b]	0.67	Messner, Berger & Nappier (2014)
Approx. BP	0.26	0.42	–	0.38	Haas et al. (1993)
Exact BP	0.167	0.191	–	0.47	Teunis & Havelaar (2000)
Exp.	–	–	0.417 2	–	Crabtree et al. (1997)
Exp.	–	–	0.019 8	–	Regli et al. (1991)
–	–	–	0.004 2	–	Haas et al. (1996)[c]
–	–	–	0.005 3	–	Teunis, Chappell & Okhuysen (2002)[c]
–	–	–	0.057 3	–	Teunis, Chappell & Okhuysen (2002)
–	0.14	1.91	–	0.070	Teunis, Chappell & Okhuysen (2002)
–	0.115	0.176	–	0.40	Teunis, Chappell & Okhuysen (2002)
–	8.37×10^{-11}	2.62×10^{-11}	–	0.76	Schijven et al. (2014)

Since the publication of infectivity results for the TAMU and UCP isolates, some risk assessment studies have used the parameter values for the hypergeometric model fitted to the combined (Iowa, TAMU and UCP) data sets (Åström et al., 2007; Smeets et al., 2007; Hunter, Zmirou-Navier & Hartemann, 2009) in order to reflect the higher predicted infectivity of some *Cryptosporidium* isolates in comparison with the model fit to the Iowa data set. Other studies (e.g. Bastos et al., 2008; Diallo et al., 2008; Carlander, Schönning & Stenström, 2009) have continued to use the parameter value for the exponential model fitted to the Iowa data set (r = 0.004). The concern is that considering only the less infectious Iowa isolate may lead to an underestimate of infectivity for environmental strains of *Cryptosporidium*. Conversely, it may be that the approximately 2 $\log_{10}$ increase in infectivity may be too conservative.

More recently, challenge study results have been published for *Cryptosporidium hominis* (Chappel et al., 2006). Schijven et al. (2014) published Beta-Poisson parameter values fitted to these data, which are included in Table D.1. The expected value of the beta distribution (r = 0.76) is higher than for previous models fitted to *C. parvum* isolates. Challenge studies have also been conducted on *C. meleagridis* (n = 5) (Chappel et al., 2011) and *C. muris* (n = 6) (Chappel et al., 2015) at high doses (1×10^5 oocysts). All subjects became infected.

D4 References

- Åström J, Petterson S, Bergstedt O, Pettersson TJ, Stenström TA (2007). Evaluation of the microbial risk reduction due to selective closure of the raw water intake before drinking water treatment. J Water Health. 5(Suppl 1):81–97.
- Atmar RL, Opekun AR, Gilger MA, Estes MK, Crawford SE, Neill FH et al. (2013). Determination of the human infectious dose-50% for Norwalk virus. J Infect Dis. 209(7):1016–22.
- Bastos RKX, Bevilacqua PD, Silva CAB, Silva CV (2008). Wastewater irrigation of salad crops: further evidence for the evaluation of the WHO guidelines. Water Sci Technol. 57(8):1213–9.
- Black RE, Levine MM, Clements ML, Hughes TP, Blaser MJ (1988). Experimental *Campylobacter jejuni* infection in humans. J Infect Dis. 157:472–9.
- Carlander A, Schönning C, Stenström TA (2009). Energy forest irrigated with wastewater: a comparative microbial risk assessment. J Water Health. 7(3):413–33.
- Chappel CL, Okhuysen PC, Langer-Curry R, Widmer G, Akiyoshi DE, Tanriverdi S et al. (2006). *Cryptosporidium hominis*: experimental challenge of healthy adults. Am J Trop Med Hyg. 75(5):851–7.
- Chappel CL, Okhuysen PC, Langer-Curry R, Akiyoshi DE, Widmer G, Tzipori S (2011). *Cryptosporidium meleagridis*: infectivity in healthy adult volunteers. Am J Trop Med Hyg. 85(2):238–42.
- Chappel CL, Okhuysen PC, Langer-Curry R, Lupo PJ, Widmer G, Tzipori S (2015). *Cryptosporidium muris*: infectivity and illness in healthy adult volunteers. Am J Trop Med Hyg. 92(1):50–5.
- Couch R, Cate T, Douglas RG, Gerone P, Knight V (1966). Effect of route of inoculation on experimental respiratory viral disease in volunteers and evidence for airborne transmission. Bacteriol Rev. 30(3):517–29.
- Crabtree KD, Gerba CP, Rose JB, Haas CN (1997). Waterborne adenovirus – A risk assessment. Water Sci Technol. 35(11–12):1–6.
- Diallo MB, Anceno AJ, Tawatsupa B, Houpt ER, Wangsuphachart V, Shipin OV (2008). Infection risk assessment of diarrhea-related pathogens in a tropical canal network. Sci Total Environ. 407(1):223–32.
- DuPont HL, Chappell CL, Sterling CR, Okhuysen PC, Rose JB, Jakubowski W (1995). The infectivity of *Cryptosporidium parvum* in healthy volunteers. N Engl J Med. 332(13):855–9.
- Evans MR, Roberts RJ, Ribeiro CD, Gardner D, Kembrey D (1996). A milk-borne *Campylobacter* outbreak following an educational farm visit. Epidemiol Infect. 117:457–62.
- Frenck R, Bernstein DI, Xia M, Huang P, Zhong W, Parker S et al. (2012). Predicting susceptibility to norovirus GII.4 by use of a challenge model involving humans. J Infect Dis. 206:1386–93.
- Gale P (2005). Land application of treated sewage sludge: quantifying pathogen risks from consumption of crops. J Appl Microbiol. 98(2):380–96.
- Haas CN (1983). Estimation of risk due to low doses of microorganisms: a comparison of alternative methodologies. Am J Epidemiol. 118(4):573–82.
- Haas CN (2002). Conditional dose–response relationships for microorganisms: development and application. Risk Anal. 22(3):455–64.
- Haas CN, Rose JB, Gerba C, Regli S (1993). Risk assessment of virus in drinking water. Risk Anal. 13:545–52.

- Haas CN, Crockett CS, Rose JB, Gerba CP, Fazil AM (1996). Assessing the risk posed by oocysts in drinking water. J Am Water Works Assoc. 88(9):131–6.
- Haas CN, Thayyar-Madabusi A, Rose JB, Gerba CP (2000). Development of a dose–response relationship for *Escherichia coli* O157 : H7. Int J Food Microbiol. 56(2–3):153–9.
- Hunter PR, Zmirou-Navier D, Hartemann P (2009). Estimating the impact on health of poor reliability of drinking water interventions in developing countries. Sci Total Environ. 407(8):2621–4.
- Mara DD, Sleigh PA, Blumenthal UJ, Carr RM (2007). Health risks in wastewater irrigation: comparing estimates from quantitative microbial risk analyses and epidemiological studies. J Water Health. 5(1):39–50.
- Masago Y, Katayama H, Hashimoto A, Hirata T, Ohgaki S (2002). Assessment of risk of infection due to *Cryptosporidium parvum* in drinking water. Water Sci Technol. 46(11–12):319–24.
- Medema GJ, Teunis PF, Havelaar AH, Haas CN (1996). Assessment of the dose–response relationship of *Campylobacter jejuni*. Int J Food Microbiol. 30(1–2):101–11.
- Medema GJ, Hoogenboezem W, van der Veer AJ, Ketelaars HAM, Hijnen WAM, Nobel PJ (2003). Quantitative risk assessment of *Cryptosporidium* in surface water treatment systems. Water Sci Technol. 47(3):241–7.
- Medema G, Loret J-C, Stenström TA, Ashbolt N (2006). Quantitative microbial risk assessment in the water safety plan. Final report on the EU MicroRisk Project. Brussels: European Commission.
- Mena KD, Gerba CP, Haas CN, Rose JB (2003). Risk assessment of waterborne coxsackievirus. J Am Water Works Assoc. 95(7):122–31.
- Messner MJ, Berger P, Nappier SP (2014). Fractional Poisson – A simple dose–response model for human norovirus. Risk Anal. 34(10):1820–9.
- Oesterholt F, Martijnse G, Medema G, van der Kooij D (2007). Health risk assessment of non-potable domestic water supplies in the Netherlands. J Water Supply Res Technol – Aqua. 56(3):171–9.
- Okhuysen PC, Chappell CL, Crabb JH, Sterling CR, DuPont HL (1999). Virulence of three distinct *Cryptosporidium parvum* isolates for healthy adults. J Infect Dis. 180(4):1275–81.
- Okhuysen PC, Rich SM, Chappell CL, Grimes KA, Widmer G, Feng X et al. (2002). Infectivity of a *Cryptosporidium parvum* isolate of cervine origin for healthy adults and interferon-knockout mice. J Infect Dis. 185:1320–5.
- Ottoson J, Stenström TA (2003). Faecal contamination of greywater and associated microbial risks. Water Res. 37(3):645–55.
- Pai CH, Kelly JK, Meyers GL (1986). Experimental infection of infant rabbits with verotoxin-producing *E. coli.* Infect Immun. 51(1):16–23.
- Powell M, Ebel E, Schlosser W, Walderhaug M, Kause J (2000). Dose–response envelope for *Escherichia coli* O157:H7. Quant Microbiol. 2:141–63.
- Regli S, Rose JB, Haas CN, Gerba CP (1991). Modeling the risk from *Giardia* and viruses in drinking water. J Am Water Works Assoc. 83(11):76–84.
- Rendtorff RC (1954). The experimental transmission of human intestinal protozoan parasites. II. *Giardia lamblia* cysts given in capsules. Am J Hyg. 59:209–20.
- Schiff GM, Stefanovic GM, Young EC, Sander DS, Pennekamp JK, Ward RL (1984). Studies of echovirus-12 in volunteers: determination of minimal infectious dose and the effect of previous infection and infectious dose. J Infect Dis. 150:858–66.
- Schijven J, Rutjes S, Smeets P, Teunis P (2014). QMRAspot: a tool for quantitative microbial risk assessment for drinking water. Manual QMRAspot version 2.0. Bilthoven: National Institute for Public Health and the Environment (RIVM Report 2014-0020).
- Seitz SR, Leon JS, Schwab KJ, Lyon GM, Dowd M, McDaniels M et al. (2011). Norovirus infectivity in humans and persistence in water. Appl Environ Microbiol. 77:6884–8.
- Shinagawa K, Hu DY, Yoshida S (1997). Correspondence and problem for enterohemorrhagic *E. coli* O157 outbreak in Morioka city, Iwate. Koshu Eisei Kenkyu. 46:104–12.
- Smeets PW, van Dijk JC, Stanfield G, Rietveld LC, Medema GJ (2007). How can the UK statutory *Cryptosporidium* monitoring be used for quantitative risk assessment of *Cryptosporidium* in drinking water? J Water Health. 5(Suppl 1):107–18.
- Smeets PW, Dullemont YJ, Van Gelder PH, Van Dijk JC, Medema GJ (2008). Improved methods for modelling drinking water treatment in quantitative microbial risk assessment; a case study of *Campylobacter* reduction by filtration and ozonation. J Water Health. 6(3):301–14.
- Soller JA, Bartrand T, Ashbolt NJ, Ravenscroft J, Wade TJ (2010). Estimating the primary etiologic agents in recreational freshwaters impacted by human sources of faecal contamination. Water Res. 44(16):4736–47.
- Suptel EA (1963). Pathogenesis of experimental coxsackie virus infection. Arch Virol. 7:61–6.

- Teunis PF, Havelaar AH (2000). The Beta Poisson dose–response model is not a single-hit model. Risk Anal. 20(4):513–20.
- Teunis PFM, Chappell CL, Okhuysen PC (2002). *Cryptosporidium* dose response studies: variation between isolates. Risk Anal. 22(1):175–83.
- Teunis PFM, Nagelkerke NJD, Haas CN (1999). Dose response models for infectious gastroenteritis. Risk Anal. 19(6):1251–60.
- Teunis PFM, Ogden ID, Strachan NJC (2008). Hierarchical dose response of *E. coli* O157:H7 from human outbreaks incorporating heterogeneity in exposure. Epidemiol Infect. 136:761–70.
- Teunis P, Takumi K, Shinagawa K (2004). Dose response for infection by *Escherichia coli* O157:H7 from outbreak data. Risk Anal. 24(2):401–7.
- Teunis PFM, van der Heijden OG, van der Giessen JWB, Havelaar AH (1996). The dose response relation in human volunteers for gastro-intestinal pathogens. Bilthoven: National Institute of Public Health and the Environment (RIVM).
- Teunis PFM, Medema GJ, Kruidenier L, Havelaar AH (1997). Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. Water Res. 31(6):1333–46.
- Teunis P, van den Brandhof W, Nauta M, Wagenaar J, van den Kerkhof H, van Pelt W (2005). A reconsideration of *Campylobacter* dose–response relation. Epidemiol Infect. 133:583–92.
- Teunis PFM, Moe CL, Liu P, Miller SE, Lindesmith L, Barie RS et al. (2008). Norwalk virus: how infectious is it? J Med Virol. 80:1468–76.
- Teunis P, Kasuga R, Fazil A, Ogden ID, Rotariu O, Strachan NJ (2010). Dose–response modeling of *Salmonella* using outbreak data. Int J Food Microbiol. 144(2):243–9.
- van den Brandhof, Wagenaar JA, van den Kerkhof H (2003). An outbreak of campylobacteriosis after drinking unpasteurized milk 2002, The Netherlands. Int J Med Microbiol. 293:548–9.
- van Heerden J, Ehlers MM, van Zyl WB, Grabow WO (2004). Prevalence of human adenoviruses in raw and treated water. Water Sci Technol. 50(1):39–43.
- van Lieverloo JH, Blokker EJ, Medema G (2007). Quantitative microbial risk assessment of distributed drinking water using faecal indicator incidence and concentrations. J Water Health. 5(Suppl 1):131–49.
- Ward RL, Bernstein DI, Young EC (1986). Human rotavirus studies in volunteers: determination of infectious dose and serological response to infection. J Infect Dis. 154(5):871–80.
- Westrell T, Schönning C, Stenström TA, Ashbolt NJ (2004). QMRA (quantitative microbial risk assessment) and HACCP (hazard analysis and critical control points) for management of pathogens in wastewater and sewage sludge treatment and reuse. Water Sci Technol. 50(2):23–30.
- WHO (in preparation). Guidelines for drinking-water quality, fourth edition incorporating first addendum. Geneva: World Health Organization.

ANNEX E | GLOSSARY

- **Akaike information criterion (AIC)** – A criterion used to select the best statistical model from a set of plausible models. One model is better than another model if it has a smaller AIC value.
- **Arithmetic mean** – The sum of the values of all samples divided by the number of samples.
- **Attack rate** – The proportion of an exposed population at risk that becomes infected or develops clinical illness during a defined period.
- **Bayesian inference** – A method of statistical inference, relying on Bayes theorem, that allows the incorporation of prior beliefs to inform the outcome of the analysis.
- **Beta distribution** – A continuous probability distribution defined on the interval [0, 1] and defined by two parameters (α and β).
- **Beta-Poisson model** – A dose–response model where the distribution in the exposure medium is assumed to be random (Poisson) and the probability of each individual organism causing infection is described by a beta distribution.
- **Conditional dose–response** – A dose–response model where the number of organisms ingested is specifically quantified (as opposed to the average dose) and the probability of infection is calculated for each exposure scenario.
- **Confidence interval** – A range of values inferred or believed to enclose the actual or true value of an uncertain quantity with a specified degree of probability. Confidence intervals may be inferred based upon sampling distributions for a statistic.
- **Control measure** – Any action or activity (or barrier) that can be used to prevent or eliminate a hazard or reduce it to an acceptable level.
- **Cost–benefit analysis** – An analysis of all the costs of a project and all the benefits. Projects that provide the most benefits at the least cost are the most desirable.
- **Critical limits** – A criterion that separates acceptability from unacceptability.
- **Cumulative density function** – For a variable x, the probability that the variable takes a value less than or equal to x.
- **Cyst** – Environmentally resistant infective parasitic life stage (e.g. *Giardia*).
- **Deterministic** – A QMRA that is undertaken with point estimates of each input value.
- **Diarrhoea** – Loose watery and frequent bowel movements, often associated with an infection.
- **Disability-adjusted life year (DALY)** – Population metric of life years lost to disease due to both morbidity and mortality.
- **Disease** – Symptoms of illness in a host (e.g. diarrhoea, fever, vomiting, blood in urine).
- **Dose** – The amount of a pathogen that enters or interacts with an organism.
- **Dose–response assessment** – The determination of the relationship between the magnitude of exposure (dose) to a chemical, biological or physical agent and the severity and/or frequency of associated adverse health effects (response).
- **Dynamic risk model** – A risk model that takes into consideration the impact of secondary spread.

- **Epidemiology** – The study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems.
- ***Escherichia coli* (*E. coli*)** – A bacterium found in the gut, used as an indicator of faecal contamination of water.
- **Excreta** – Faeces and urine.
- **Expert judgement** – A reasoned formation of opinions by someone with special knowledge or experience in a particular problem domain. Expert judgement is documented and can be explained to satisfy outside scrutiny.
- **Exposure** – Contact of a chemical, physical or biological agent with the outer boundary of an organism (e.g. through inhalation, ingestion or dermal contact).
- **Exposure assessment** – The estimation (qualitative or quantitative) of the magnitude, frequency, duration, route and extent of exposure to one or more contaminated media.
- **Faecal indicator bacteria** – Bacteria used to detect and estimate the level of faecal contamination in water.
- **Faecal indicator organisms** – Microorganisms used to detect and estimate the level of faecal contamination in water (includes faecal indicator bacteria).
- **Faecal sludge** – Sludges of variable consistency collected from on-site sanitation systems, such as latrines, non-sewered public toilets, septic tanks and aqua privies. Septage, the faecal sludge collected from septic tanks, is included in this term.
- **First-order decay –** When the change in microbial concentration over time on a $\log_{10}$ scale is linear and can be described by one parameter.
- **Gamma distribution** – A two-parameter continuous probability distribution with domain >0. When concentration is gamma distributed, the resulting distribution of microbial counts follows a negative binomial distribution.
- **Geometric mean** – A measure of central tendency, just like a median. It is different from the traditional mean (which is called the arithmetic mean) because it uses multiplication rather than addition to summarize data values. The geometric mean is a useful summary when changes in the data occur in a relative fashion.
- **Greywater** – Water from the kitchen, bath and/or laundry, which generally does not contain significant concentrations of excreta.
- **Groundwater** – Water contained in rocks or subsoil.
- **Hazard** – A biological, chemical, physical or radiological agent that has the potential to cause harm.
- **Hazardous event** – An event in which people are exposed to a hazard within the system. It may be an incident or a situation that introduces or releases the hazard to the environment in which humans are living or working; amplifies the concentration of a hazard; or fails to remove a hazard from the human environment.
- **Health-based target** – A defined level of health protection for a given exposure. This can be based on a measure of disease (e.g. 10^{-6} DALY per person per year) or the absence of a specific disease related to that exposure.
- **Health effects assessment** – The process of evaluating and quantifying the health impact of reference pathogens in a QMRA and may include dose–response, probability of illness and other adverse health outcomes and DALYs.
- **Human infectivity** – Capable of causing infection in a human host.
- **Illness** – A condition marked by pronounced deviation from the normal healthy state.
- **Immunity** – Inability of an individual to become infected due to protection via his or her immune system.
- **Indicator organisms** – Microorganisms whose presence is indicative of faecal contamination and possibly of the presence of more harmful microorganisms.
- **Infection** – The entry and development or multiplication of an infectious agent in a host. Infection may or may not lead to disease symptoms (e.g. diarrhoea). Infection can be measured by detecting infectious agents in excreta or colonized areas or through measurement of host immune response (i.e. the presence of antibodies against the infectious agent).
- **Infectivity –** Capable of causing infection in a host.
- **Likelihood** – The probability that the observed data were generated from an assumed model.
- **Lognormal distribution –** A two-parameter continuous probability distribution with domain >0.
- **Log reduction** – Organism removal efficiencies: 1 $\log_{10}$ unit = 90%; 2 $\log_{10}$ units = 99%; 3 $\log_{10}$ units = 99.9%; and so on.
- **Markov chain Monte Carlo** – A general method of sampling arbitrary high-dimensional probability distributions by taking a random walk through configuration space. Can be used to obtain a distribution of the likely values of the parameter(s) of interest given the observed data.
- **Median** – The middle value of a sample series (50% of the values in the sample are lower and 50% are higher than the median).

- **Meta-analysis** – A method for systematically combining quantitative data from several studies to develop a single conclusion or parameter estimate.
- **Model** – A set of constraints restricting the possible joint values of several quantities; a hypothesis or system of beliefs regarding how a system works or responds to changes in its inputs. The purpose of a model is to represent a particular system of interest as accurately and precisely as necessary with respect to the particular decision objectives.
- **Model structure** – A set of assumptions and inference options upon which a model is based, including underlying theory as well as specified functional relationships.
- **Model uncertainty** – Uncertainty associated with the selection of models (including probability distributions and process models) for describing environmental and biological processes in the QMRA.
- **Monte Carlo simulation** – A method of quantifying the distribution of an outcome (i.e. exposure or risk) by random simulation from the probability distributions describing all model input variables.
- **Most probable number** – The maximum likelihood estimator (MLE) of the mean concentration based on the microbial presence/absence results, and assuming a Poisson distribution of organisms.
- **Multiple barriers** – Use of more than one preventive measure as a barrier against hazards.
- **Negative binomial distribution** – A discrete two-parameter distribution that can be used to describe microbial counts. When the concentration is assumed to vary according to a gamma distribution, the resulting counts will follow a negative binomial distribution.
- **Non-culturable** – Unable to be amplified by culture in the laboratory.
- **Oocyst** – A structure that is produced by some coccidian protozoa (e.g. *Cryptosporidium*) as a result of sexual reproduction during the life cycle. The oocyst is usually the infectious and environmental stage, and it contains sporozoites. For the enteric protozoa, the oocyst is excreted in faeces.
- **Operational target** – Operational objectives or performance goals established to indicate the performance of a control measure. Deviations from operational limits should be considered a trend towards loss of control of a process. Operational limits are typically more stringent than critical limits, to allow timely corrective action to be taken before a critical limit is breached.
- **Parameter** – A quantity used to calibrate or specify a model, such as parameters of a probability model (e.g. mean and standard deviation for a normal distribution). Parameter values are often selected by fitting a model to a calibration data set.
- **Parameter uncertainty** – Uncertainty associated with the estimated value of model parameters given an assumed model and the observational data.
- **Parametric distribution** – A probability distribution defined by parameters.
- **Pathogen** – A disease-causing organism (e.g. bacteria, helminths, protozoa and viruses).
- **Point estimate** – A single quantitative value used to define a model variable.
- **Poisson distribution** – A probability distribution that describes a discrete random variable (i.e. variables that may take on only a countable number of distinct values, such as 0, 1, 2, 3, 4, ...).
- **Polymerase chain reaction** – A relatively simple enzymatic reaction used to generate copies of a target DNA sequence through a series of temperature cycles.
- **Precision** – A measure of the reproducibility of the predictions of a model or repeated measurements, usually in terms of the standard deviation or other measures of variation among such predictions or measurements.
- **Probabilistic analysis** – Analysis in which distributions are assigned to represent variability or uncertainty in quantities. The form of the output of a probabilistic analysis is likewise a distribution.
- **Probability** – Degree of belief regarding the likelihood of a particular range or category (e.g. the probability that an average individual with a particular mean dose will develop an illness).
- **Probability density function** – A curve that specifies by means of the area under the curve over an interval the probability that a continuous random variable falls within the interval.
- **Probability distribution** – A function that, for each possible value of a discrete random variable, takes on the probability of that value occurring.
- **Problem formulation** – A systematic planning step that identifies the goals, breadth and focus of the risk assessment, the regulatory and policy context of the assessment and the major factors that will need to be addressed for the assessment.
- **Quantitative microbial risk assessment** – Microbial risk assessment when each component in the model is specifically quantified.

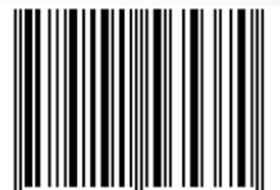
- **Reference pathogen** – Pathogen selected to represent a broader group of pathogens. If the system is designed to protect public health from the reference pathogen, then it is assumed that public health will be protected from all pathogens in the broader group.
- **Representativeness** – The extent to which a sample represents the magnitude and variability of the value of interest.
- **Risk** – The likelihood that a hazardous event occurs, and the severity or consequence of the hazard.
- **Risk assessment** – The overall process of using available information to predict how often specified events may occur (likelihood) and the magnitude of their consequences.
- **Risk management** – The process, distinct from risk assessment, of weighing policy alternatives, in consultation with all interested parties, considering risk assessment and other factors relevant for the health protection of consumers and promotion of fair trade practices, and, if needed, selecting appropriate prevention and control options.
- **Sanitary inspection** – An on-site inspection and evaluation by qualified individuals of all conditions, devices and practices in the water supply system that pose an actual or potential danger to the health and well-being of the consumer. Sanitary inspection typically makes use of standardized "sanitary inspection forms".
- **Sanitation safety plan** – A step-by-step risk-based approach to assist in the implementation of the WHO guidelines for the safe use of wastewater, excreta and greywater in agriculture and aquaculture.
- **Scenario** – A specific situation or circumstance that is modelled within the QMRA.
- **Schmutzdecke** – Complex biological scum layer formed at the top of a filter; essentially the main treatment layer of the filter.
- **Secondary transmission** – The direct or indirect propagation of a pathogen from an infected person (with or without clinical illness) to additional people.
- **Sensitivity analysis** – A method used to examine the behaviour of a model by measuring the variation in its outputs resulting from changes in its inputs.
- **Sequelae** – Severe, secondary and/or chronic health effects that occur following initial infection.
- **Sewage** – Mixture of human excreta and water used to flush the excreta from the toilet and through the pipes; may also contain water used for domestic purposes.
- **Single-hit model** – A dose–response model that relies on the assumption that every microorganism to which an individual is exposed has an equal and independent chance of causing infection.
- **Statistic** – A function of a random sample of data (e.g. mean, standard deviation, distribution parameters).
- **Statistical distribution** – A probability distribution.
- **Statistical inference** – A systematic approach used to draw conclusions from data sets arising from systems affected by random variation.
- **Stochastic** – An analysis approach that allows the quantification of the probability distribution of risk; typically undertaken using Monte Carlo simulation.
- **Surface water** – All water naturally open to the atmosphere (e.g. rivers, streams, lakes and reservoirs).
- **Surrogate** – A parameter used to represent or model the behaviour of the pathogen of interest.
- **Surrogate data** – Substitute data or measurements on one quantity used to estimate analogous or corresponding values for another quantity.
- **Thermotolerant coliforms** – Group of bacteria whose presence in the environment usually indicates faecal contamination; previously called faecal coliforms.
- **Tolerable health risk** – Defined level of health risk from a specific exposure or disease that is tolerated by society; used to set health-based targets.
- **Transparent** – Characteristics of a process where the rationale, logic of development, constraints, assumptions, value judgements, limitations and uncertainties of the expressed determination are fully and systematically stated, documented and accessible for review.
- **Uncertainty** – Lack of knowledge regarding the true value of a quantity, such as a specific characteristic (e.g. mean variance) of a distribution for variability, or regarding the appropriate and adequate inference options to use to structure a model or scenario (also referred to as *model uncertainty* and *scenario uncertainty*). Uncertainty can be reduced by obtaining more information through research and data collection, such as through research on mechanisms, larger sample sizes or more representative samples.
- **Validation** – Testing the system and its individual components to prove that it is capable of meeting the specified targets (i.e. microbial reduction targets); should take place when a new system is developed or new processes are added.

- **Variability** – Observed differences attributable to true heterogeneity or diversity in a population or exposure parameter. Variability implies real differences among members of that population. For example, different individuals have different intakes and susceptibilities.
- **Viability** – A measure that indicates if a microorganism has some characteristic or characteristics that are consistent with infectivity.
- **Wastewater** – Liquid waste discharged from homes, commercial premises and similar sources to individual disposal systems or to municipal sewer pipes, and which contains mainly human excreta and used water. When produced mainly by household and commercial activities, it is called domestic or municipal wastewater or domestic sewage. In this context, domestic sewage does not contain industrial effluents at levels that could pose threats to the functioning of the sewerage system, the treatment plant, public health or the environment.
- **Water safety plan** – A comprehensive risk assessment and management approach to ensuring drinking-water safety that encompasses all steps in the water supply from catchment to consumer.

Sources of definitions:

- Assessing human health risks of chemicals: derivation of guidance values for health-based exposure limits. Geneva: World Health Organization, International Programme on Chemical Safety; 1994 (Environmental Health Criteria 170; http://www.inchem.org/documents/ehc/ehc/ehc170.htm, accessed 12 April 2016).
- Bartram J, Corrales L, Davison A, Deere D, Drury D, Gordon B et al. Water safety plan manual: step-by-step risk management for drinking-water suppliers. Geneva: World Health Organization; 2009 (http://www.who.int/water_sanitation_health/publications/publication_9789241562638/en/, accessed 12 April 2016).
- Dorland's illustrated medical dictionary, twenty-sixth edition. Philadelphia (PA): W.B. Saunders Company; 1981.
- Guidelines for drinking-water quality, second edition. Vol. 3. Surveillance and control of community supplies. Geneva: World Health Organization; 1997 (http://www.who.int/water_sanitation_health/publications/gdwq3rd_2ndadd/en/, accessed 28 May 2016).
- Guidelines for the safe use of wastewater, excreta and greywater. Vol. 4. Excreta and greywater use in agriculture (http://www.who.int/water_sanitation_health/publications/gsuweg4/en/, accessed 6 May 2016).
- Hazard characterization for pathogens in food and water. Guidelines. Rome: Food and Agriculture Organization of the United Nations; and Geneva: World Health Organization; 2003 (Microbiological Risk Assessment Series, No. 3; http://apps.who.int/iris/bitstream/10665/42693/1/9241562374.pdf, accessed 12 April 2016).
- Last JM, editor. A dictionary of epidemiology, third edition. New York (NY): Oxford University Press; 1995.
- Meynell GG, Stocker BAD. Some hypotheses on the aetiology of fatal infections in partially resistant hosts and their application to mice challenged with *Salmonella paratyphi* – B or *Salmonella typhimurium* by intraperitoneal injection. J Gen Microbiol. 1957;16:38–58.
- Microbial risk assessment guideline: pathogenic microorganisms with focus on food and water. Prepared by the Interagency Microbiological Risk Assessment Guideline Workgroup. United States Department of Agriculture, Food Safety and Inspection Service, and United States Environmental Protection Agency; 2012 (USDA/FSIS/2012-001; EPA/100/J12/001; http://www.fsis.usda.gov/wps/wcm/connect/d79eaa29-c53a-451e-ba1c-36a76a6c6434/Microbial_Risk_Assessment_Guideline_2012-001.pdf?MOD=AJPERES, accessed 12 April 2016).
- Principles and guidelines for the conduct of microbiological risk assessment. Rome: Food and Agriculture Organization of the United Nations and World Health Organization, Joint FAO/WHO Food Standards Programme, Codex Alimentarius Commission; 1999 (Document No. CAC/GL-30; http://www.fao.org/docrep/004/y1579e/y1579e05.htm, accessed 12 April 2016).
- Procedural manual, twenty-fourth edition. Rome: Food and Agriculture Organization of the United Nations and World Health Organization, Joint FAO/WHO Food Standards Programme, Codex Alimentarius Commission; 2015 (ftp://ftp.fao.org/codex/Publications/ProcManuals/Manual_24e.pdf, accessed 6 May 2016).
- Revised framework for microbial risk assessment. An ILSI Risk Science Institute workshop report. Washington (DC): International Life Sciences Institute; 2000.
- Sanitation safety planning: manual for safe use and disposal of wastewater, greywater and excreta. Geneva: World Health Organization; 2015 (http://www.who.int/water_sanitation_health/publications/ssp-manual/en/, accessed 12 April 2016).

ISBN 978 92 4 156537 0



World Health Organization