MODELLING THE PUBLIC HEALTH RISKS ASSOCIATED WITH ENVIRONMENTAL EXPOSURES: A CASE STUDY IN WASTEWATER REUSE

Denise Anne Beaudequin

BSc - Griffith University
MEnvCH - Griffith University
MNsg(ProfSt) - Queensland University of Technology

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Science and Engineering Faculty
Queensland University of Technology
2016

Keywords

Adaptive management, Bayesian networks, causality, complex systems, exposure, probability, quantitative microbial risk assessment, uncertainty, wastewater, water recycling.

Abstract

Water recycling is an important solution to escalating global issues of water shortage. However established methods of ascertaining microbial safety of wastewater, such as quantitative microbial risk assessment (QMRA), are constrained by issues such as poor quality or paucity of data and inability to capture the many factors influencing risk in exposure pathways. In the absence of comprehensively characterised exposure scenarios, universal standards are frequently used for recycled water to minimise risk, unnecessarily increasing treatment costs and inhibiting uptake of reuse schemes. Bayesian networks (BNs) offer a systems approach to characterisation of complexity and are increasingly recognised as a powerful, flexible tool to address many of the issues with existing methods.

To develop and evaluate applications of BNs in this field, four objectives were defined. The first was to gauge the extent and nature of applications of BNs associated with QMRA and identify and fill a gap in the literature. The second objective was to create a conceptual model of key variables influencing health risk in a water recycling context. The third objective was to develop a prototype BN to assess and manage health risk in a wastewater reuse scenario. The fourth objective was to develop concurrent BNs representing key pathogen groups for water recycling and evaluate their utility in assessment and management of wastewater treatment and reuse.

Chapter 1 outlines the context, significance and scope of the research and provides an overview of the thesis. A broad review of the literature relevant to the research problem and proposed solution is then undertaken in Chapter 2. In Chapter 3, the specific literature on applications of BNs in QMRA is reviewed and analysed in detail. Chapter 4 describes a conceptual model of health risk associated with wastewater reuse based on the QMRA framework, developed from peer-reviewed works and evaluated by industry stakeholders.

In Chapter 5 development of a QMRA expressed as a BN in a context of consumption of wastewater-irrigated lettuce is described, for evaluating a range of exposure and risk mitigation scenarios. The BN revealed that lettuce washing and

withholding irrigation had more influence on infection risk than other variables, including pathogen concentration in treated water.

In Chapter 6, models representing the principal waterborne pathogen groups were developed for assessment and management of risk associated with wastewater irrigation of public open space. These models incorporated in-treatment and post-treatment risk reduction strategies and multiple exposure profiles. In a scenario involving poorly treated water, onsite risk reduction measures alone significantly increased the chance of tolerable disease burden. In another scenario, chlorination was shown to have an insignificant effect on disease burden, relative to reducing frequency of exposure. To construct the BNs described in Chapters 5 and 6, deterministic and stochastic QMRA models were developed using values from peer-reviewed literature and data was generated using Monte Carlo simulation.

Chapter 7 summarises and discusses the findings of the research. BNs offer a number of features for addressing QMRA constraints. They enable better understanding of complex scenarios through the graphic portrayal of risk pathways, the quantification of variables for which there may be little or no data and the explicit representation of knowledge limitations and uncertainty in the studied system. The advantages of BNs include an accessible visual platform, the ability to quantify relationships between variables and the use of probability distributions to represent uncertainty. BNs are capable of predictive and scenario analysis with instant updating and thus facilitate adaptive management. The drawbacks of using BNs include their inability to support feedback loops, elicitation of the conditional probabilities, loss of information as a result of discretising continuous variables and assumptions regarding prior distributions.

To make this research accessible to and useful for industry stakeholders, a plain-language summary of the rationale for and procedures underlying the BN methodology has been included as Appendix A.

This work represents a novel approach to modelling microbial risk, employing recently-developed statistical methodology for the first time to quantify microbial risk associated with wastewater reuse. By utilising the features of BNs, multiple objectives identified in the literature have been fulfilled: the BNs portray and quantify complex exposure-health relationships; incorporate risk assessment and management options for wastewater reuse scenarios; employ the multiple barrier

approach to risk management; enable integration of traditional microbial indicators with health outcome targets to limit disease; and facilitate the adaptive management paradigm. In the assessment and management of health risk related to water reuse, BNs provide a transparent, defensible evidence base for water resource managers, operators and engineers, regulatory authorities, risk modellers and water scientists to describe and quantify risk pathways, compare decision options and predict outcomes of management policies. This research clearly establishes the significant utility and potential of BN modelling for characterisation of microbial risk and validates QMRA-based BNs as an accessible tool to facilitate fit-for-purpose water recycling.

Table of Contents

List of Figures

List of Tables

List of Publications

Beyond QMRA: Modelling microbial health risk as a complex system using Bayesian networks	Beaudequin, D., Harden, F., Roiko, A., Stratton, H., Lemckert, C., & Mengersen, K. (2015). Beyond QMRA: Modelling microbial health risk as a complex system using Bayesian networks. *Environment International, 80*, 8-18.
Modelling microbial health risk of wastewater reuse: A systems perspective	Beaudequin, D., Harden, F., Roiko, A., Stratton, H., Lemckert, C., & Mengersen, K. (2015). Modelling microbial health risk of wastewater reuse: A systems perspective. *Environment International, 84*, 131-141.
Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water	Beaudequin, D., Harden, F., Roiko, A., & Mengersen, K. (2016). Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water. *Science of the Total Environment, 541*, 1393–1409.
Potential of Bayesian networks for adaptive management in water recycling	Beaudequin, D., Harden, F., Roiko, A., & Mengersen, K. (Submitted). *Environmental Modelling and Software.*

We the co-authors agree to the use of the papers in this thesis in fulfilment of the requirements for the degree of Doctor of Philosophy:

QUT Verified Signature

Kerrie Mengersen — 14/3/16 date

QUT Verified Signature

Fiona Harden — 18.01.16 date

QUT Verified Signature

Anne Roiko — 18/01/2016 date

QUT Verified Signature

Helen Stratton — 21/01/2016 date

QUT Verified Signature

Charles Lemckert — 22/1/16 date

List of Abbreviations

BN	Bayesian network
BOD	biochemical oxygen demand
BOM	Bureau of Meteorology
CAMRA	Center for Advancing Microbial Risk Assessment
CFU	colony forming unit
DAF QLD	Department of Agriculture and Fisheries Queensland
DAG	directed acyclic graph
DALY	disability-adjusted life years
EPA QLD	Environmental Protection Agency Queensland
FAO/WHO	Food and Agriculture Organization of the United Nations/World Health Organization
FIB	faecal indicator bacteria
HIV/AIDS	human immunodeficiency virus/acquired immune deficiency syndrome
ILSI	International Life Sciences Institute
IOM	Institute of Medicine
IPCC	Intergovernmental Panel on Climate Change
ISI	Institute for Scientific Information
LRV	log removal value
MC	Monte Carlo
MPN	most probable number
MPRM	modular process risk model
NHMRC	National Health and Medical Research Council
NRC	National Research Council
NRC/CIWP	National Research Council Committee on Indicators for Waterborne Pathogens
NRMMC-EPHC-AHMC	Natural Resource Management Ministerial Council, Environment Protection and Heritage Council and Australian Health and Medical Council
NWC	National Water Commission
PCR	polymerase chain reaction
PPPY	per person per year
QMRA	quantitative microbial risk assessment

UK	United Kingdom
UN	United Nations
UP DSL	University of Pittsburgh Decision Systems Laboratory
USEPA	United States Environmental Protection Agency
USEPA-USDA/FSIS	United States Environmental Protection Agency, United States Department of Agriculture/Food Safety and Inspection Service
USFDA	United States Food and Drug Administration
WHO	World Health Organisation
WRA	Water Research Australia

Glossary of Terms

algorithm	a mathematical procedure to be followed in calculations, especially by a computer
backwards inference	a useful property of a Bayesian network which enables discovery of conditions required 'upstream' to achieve a desired node outcome. Also referred to as diagnostic reasoning
Bayesian network	a probabilistic, graphical model, comprising variables represented by nodes and causal relationships between the variables, represented by arrows
causality	the relationship between a variable and the factors influencing it; indicated by an arrow in a BN. The node at the head of the arrow is influenced by the node at the tail of the arrow
chance	the measure of the likelihood that an event will occur, expressed as a percent, quantified by a number between 0% (impossibility) and 100% (certainty)
chance node	a variable represented by a probability distribution of its states (e.g., high = 0.1, medium = 0.7, low = 0.2)
child node	node with influencing factors indicated by incoming arrows from other nodes
conditional probability	probability of an event that is dependent upon another event
conditional probability table	a table underlying a child node containing the conditional probabilities for all possible combinations of influencing node states
deterministic	a deterministic model is one in which inputs are point estimates and which given the same input information will always produce the same output information
dichotomous	a case of discretisation in which the number of discrete classes is two
discretise	the process of converting continuous data to discrete categories or 'states' e.g., high or low, using chosen threshold values
downstream	at a subsequent point in a Bayesian network, closer to target nodes, in the direction of the arrows

forwards inference	ability of a Bayesian network to support 'what if' analysis by determining the effect of changes in upstream variables on target nodes. Also referred to as predictive reasoning
hyperparameter	In Bayesian statistics, a hyperparameter is a parameter of a prior distribution; the term is used to distinguish these from parameters of the model for the underlying system under analysis.
illness/disease	signs and symptoms of infection in a host
infection	invasion and multiplication by a microorganism in a host, as defined by a clinical indication such as antibody rise in the blood; may or may not be accompanied by signs and symptoms of illness in the host e.g., rash, fever, sore throat
introduction of new evidence in a BN	in a chance node, this means setting a node to 100% certainty for one of its states (or reversing that change)
irrigation withholding period	a period of time between time of last irrigation with recycled water and time of potential exposure, (e.g. lettuce harvest or public access to a park), introduced to allow microbial die-off to occur
joint distribution	the mutual distribution of all states in all nodes in a Bayesian network, taking into account node dependencies and any new evidence introduced to the network. The joint distribution is calculated by software algorithms and is expressed in individual nodes as a probability distribution across the node's states
model	representation (verb or noun) of an entity, a process or a system; can be mathematical, graphical or conceptual
Monte Carlo simulation	simulation by repeated random sampling to obtain numerical results
node	in a BN, a node represents a variable or unknown quantity
parameter	a characteristic, feature or measurable factor; including variables and constants
posterior	beliefs or probability distributions in a Bayesian network after new evidence is introduced and the network updated
priors	beliefs or probability distributions in a Bayesian network before new evidence is introduced and the network updated
probabilistic	based on probability

probability	the measure of the likelihood that an event will occur, quantified by a number between 0 (impossibility) and 1 (certainty)
quantitative microbial risk assessment	a structured approach which brings information and data together with mathematical models to examine the exposure and spread of microbial agents and to characterise the nature of the adverse outcomes
response node	any node in a BN that represents an outcome of interest; depends on the question being asked
risk estimate	QMRA risk estimates describe the probability of infection or illness in an individual or a population as a result of exposure to pathogens in a specific scenario e.g., 'the median annual norovirus disease burden was estimated to be 5.95×10^{-4} DALY/person/year'.
root node	node with no incoming arrows, i.e., no influencing factors
sensitivity analysis	reveals how sensitive an output is to any change in an input while keeping other inputs constant; can be achieved by varying the value of one input at a time and assessing the effect on an output, or through use of algorithms
simulation	the representation of the behaviour or characteristics of a system through the use of a mathematical model or a computer program
states	mutually exclusive categories (nominal or ordinal) in a chance variable
stochastic	a stochastic model has one or more random elements and the output is hence unpredictable
target node	any node in a BN that represents an outcome of interest; can depend on the question being asked
uncertainty	lack of perfect knowledge about a variable value, which can be reduced by further measurements sources of uncertainty in QMRA include: • the chosen model • representativeness of samples • enumeration methods, e.g., assumed pathogen/indicator relationships • dose-response equation parameters • non-differentiation between strains of microorganisms • assumptions/definitions, e.g., viable but non-

	culturable organisms may still cause infection
variability	the spread of a set of measurements of a variable that is a consequence of the physical system (i.e., individual or environmental variability) and that cannot be reduced by additional measurements
variable	a characteristic, feature or measurable factor that is likely to change (e.g., pathogen concentration)

Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature:

QUT Verified Signature

Date: 13 October 2016

Dedication

To Dominic and Derek; each equally my pride and joy.

Acknowledgements

I am forever indebted to my QUT supervisors, Professor Kerrie Mengersen and Dr Fiona Harden, for their ready encouragement and guidance throughout this process. I could not have imagined a more highly-principled, effective and inspirational team. I am profoundly grateful to Kerrie for passing on some of her formidable statistical knowledge. Despite her arduous schedule, she always managed to respond from wherever she was in the world. My heartfelt thanks go to Fiona for sharing her exceptional gift for language, for her swift replies to any communication and for her commitment to her students.

Sincere thanks go to my external supervisor, Associate Professor Anne Roiko, who encouraged me to pursue a PhD and suggested a project that was a perfect match for my experience and interests. I would also like to acknowledge the support of other members of the 'Pond Project' team, particularly Associate Professor Helen Stratton, Professor Charles Lemckert, Dr John Xie, Dr Edoardo Bertone, Katrina Kelly and Sonya Kozak. Thank you also to my local external supervisor, Associate Professor Peter Dunn.

I would like to express my gratitude to members of my seminar panel, Adjunct Associate Professor Dr Jim Smith and Dr Paul Wu; their contributions were invaluable.

A special acknowledgement goes to my secondary school biology teacher Eileen Brown SGS, who kindled my fascination for biological and environmental sciences, and who was a wonderful role model for women in science.

Thank you to professional editor Robyn Kent who provided assistance with formatting.

Chapter 1: Introduction

1.1 PROBLEM DESCRIPTION

The characterisation of the human health impacts of environmental exposure to pathogens is complex and challenging. Quantitative microbial risk assessment (QMRA), a structured approach to the assessment of health risks from pathogenic organisms in food and water, uses mathematical models to examine the exposure and spread of microbial agents and characterise the nature of adverse outcomes (Haas et al., 2014, USEPA-USDA/FSIS, 2012). QMRA however, is inevitably dependent upon quantitative data for model execution and realisation of conclusions, and dependable data to populate QMRA models is often difficult to obtain. Due to the microscopic nature of the subject, enumeration of microorganisms can be challenging, costly and not always achievable (O'Toole, 2011, O'Toole et al., 2008). In the characterisation of microbial exposures, there is a multiplicity of exposure routes, frequencies, media and temporal and spatial variability to consider. Widespread uncertainty can result from the choice of model, differential data quality and reliability due to disparate enumeration methods, variability in the environmental system and the variance in the estimates produced. The breadth and variability of the environmental domain also often equates to knowledge gaps where data do not exist (Haas, 2002). In a context of water recycling, assessing and managing exposures to microbial hazards under uncertain conditions is challenging for decision makers. Water utilities managers, treatment plant operators or regulatory authorities may be faced with choosing a course of action based on imperfect risk estimates, potentially resulting in unknown outcomes. Without well mapped, quantified exposure pathways, blanket standards are frequently used for recycled water to minimise risk, driving up treatment costs and inhibiting uptake of reuse schemes.

1.2 POTENTIAL SOLUTION

Bayesian networks (BNs) have been used in this study as a complementary approach to QMRA to overcome some of the limitations described. BNs are powerful integrative tools that provide probabilistic solutions to complex, causal problems and are useful for supporting decision making under uncertainty (Jensen

and Nielsen, 2007, Korb and Nicholson, 2011, Pearl, 2000). BNs offer a number of features that address the particular challenges in risk assessment and management associated with environmental exposures to microbial hazards (Parsons et al., 2005, Greiner et al., 2013). These features include the ability to study multiple interacting variables simultaneously and to accommodate missing, sparse or inaccurate data. Data of different types can be combined with expert opinion, or a BN can be constructed entirely from expert opinion. BNs can be used for causal reasoning, supporting network queries such as what-if scenarios. They can also be used for inferential reasoning, working backwards to find out which variables are key drivers for an outcome. Scenario or 'what if' analysis is efficient, because a BN responds immediately to changes such as the introduction of new evidence. As they are graphical models, BNs are represented on a clear, visual platform that promotes multidisciplinary collaboration and stakeholder engagement. Uncertainty in BN models is represented transparently at variable level, in probability distributions of variable states. The knowledge engineering cycle underlying the BN concept is an iterative process, supporting adaptive management, a constructive paradigm used in the management of complex environmental systems.

The aim of this research is therefore to develop a complex systems model of the human health risks associated with exposures to microbial pathogens in the nonpotable reuse of treated wastewater. The overarching purpose of the work is to provide a novel approach that more credibly represents microbial risks, to facilitate greater accuracy and science-based decision making with regards to fit-for-purpose wastewater treatment and reuse. As a relatively emergent technique, BNs have not been widely used in the QMRA domain and have had little previous application in assessing and managing health risk associated with wastewater reuse. This thesis represents a new approach to characterisation of microbial exposures, employing recently developed statistical methodology to portray and quantify complex exposure-health relationships. This body of work is the first instance in which the BN modelling has been used to augment QMRA in a water recycling context.

1.3 OBJECTIVES OF THE RESEARCH

The objectives of the research are:

1. To identify and fill a gap in the peer-reviewed literature on applications of BNs in QMRA (Chapter 3);

2. To develop a conceptual model of influences on microbial health risk in a wastewater reuse context (Chapter 4);

3. To develop and evaluate a BN model for the assessment and management of microbial health risk in the context of wastewater reuse (Chapter 5);

4. To develop concurrent BNs representing the principal waterborne pathogen groups for water recycling and to validate their utility in assessment and management of wastewater treatment and reuse (Chapter 6).

1.4 CONTEXT OF THE RESEARCH – WASTEWATER REUSE

The collective impacts on global water resources of population growth, increased water demands and regional water scarcities due to climate change have resulted in the worldwide increase in prominence of the practice of reclaiming and reusing wastewater, particularly in arid regions (Bitton, 2005). Recycling of waters that have previously been regarded as unusable can provide additional sources of water for a range of purposes that are unnecessarily supplied by limited freshwater resources. Moreover, use of treated wastewater in irrigation, cleaning or industry has the potential to reduce costs, energy and resource consumption through customisation of treatment requirements to provide a fit-for-purpose resource. However, efficient assessment and management of the microbiological health risks associated with waters treated to varying levels of quality for different purposes is difficult to achieve due to issues such as data scarcity, expensive or difficult assay methods and the number of exposure pathways and causal variables requiring consideration.

1.5 PURPOSE OF THE RESEARCH

Faecal indicator organism levels or pathogen concentrations alone are inadequate for judging health risk in reclaimed water exposures, as there are numerous other factors in exposure pathways contributing to the final dose to which an individual is exposed (Haas et al., 2014, NRC/CIWP, 2004). There is a

widespread need for the use of QMRA to realistically determine the microbial suitability of reclaimed water for specific uses (Soller et al., 2016, Ashbolt et al., 2010, Bichai and Smeets, 2013). This study has developed probabilistic graphical models to integrate important influential variables in potential exposure pathways. The models incorporate indications of wastewater treatment performance and other exposure variables with potential risk reduction strategies, to produce a holistic evaluation of microbial health risk.

1.6 SIGNIFICANCE OF THE RESEARCH

BNs have been used to some extent with QMRA but chiefly in the area of food risk assessment. To the author's knowledge, there have only been two uses of BNs in the wastewater and health risk area. This work will be an important addition to the seminal applications of BNs in this domain by Donald et al. (2009) and Cook et al. (2011) and will contribute a novel application of the method to health risk assessment in water recycling.

Risk assessment is not a standalone process. The established risk paradigm described by the National Research Council (NRC) describes two interlinked processes, risk assessment and risk management (NRC, 2009, NRC, 1983). While the aim of risk assessment is to evaluate the degree and probability of harm to human health from an adverse effect or event, it should be emphasised that the assessment of risk is not an objective in its own right, but forms the basis for the decision-making process of risk management. Risk assessment can be a starting point in an iterative cycle comprising risk assessment and risk management (Fewtrell et al., 2001). The purpose of risk management then, is to identify and prioritise public health or environmental risks and enact decisions in the public health interest. Such decisions need to be based on social and economic factors as well as optimal application of resources "to sustainably minimize, monitor and control the adverse impact events or to maximize the realization of opportunities" (NRC, 1983).

The assessment and management of risk in environmental systems is complex and sometimes controversial, due to inherent uncertainty and variability. The adaptive management paradigm (IOM, 2013), described elsewhere as 'learning as we go' (Laniak et al., 2013), is commonly used in management of natural resources (Chen and Pollino, 2012, Nyberg et al., 2006, Pollino and Henderson, 2010).

Adaptive management is based on an iterative decision making, monitoring and learning cycle, improving long term management outcomes through making short term decisions, observing the outcomes and modifying management strategies as understanding of the system improves (Holling, 1978, Walters, 1986). Similar to the 'plan-do-check-act' quality improvement method used in business for control and continuous improvement of processes and products (Walton and Deming, 1986), adaptive management brings about robust decision making in the face of commonly encountered uncertainty in environmental domains. Instead of using a single set of probability distributions, adaptive management strategies use multiple representations of the future, or scenarios, to characterise and reduce uncertainty (Lempert and Collins, 2007). BNs are well suited to adaptive management approaches, as they support rapid 'what if' analyses and iterative improvement methods. This thesis demonstrates the utility of BNs in incorporating risk management options, together with risk assessment variables and their capacity for efficient scenario analyses to gauge public health risk.

1.7 SCOPE AND LIMITATIONS OF THE RESEARCH

This body of work encompasses five of the six steps in the generic risk assessment/management framework: hazard identification, exposure assessment, dose response assessment, risk characterisation and risk management (NRC, 2009, NRC, 1983) (Figure 1.1). The study focuses on irrigation as a water reuse option and the microbial risk assessments underlying the BNs in this study are static risk assessments. The chosen pathogens for the QMRAs underlying the BNs in the study are norovirus, *Cryptosporidium parvum* and *Campylobacter jejuni*, representing the three major waterborne pathogen groups viruses, protozoa and bacteria (NRMMC-EPHC-AHMC, 2006). Water ingestion is the chosen exposure route for examination.

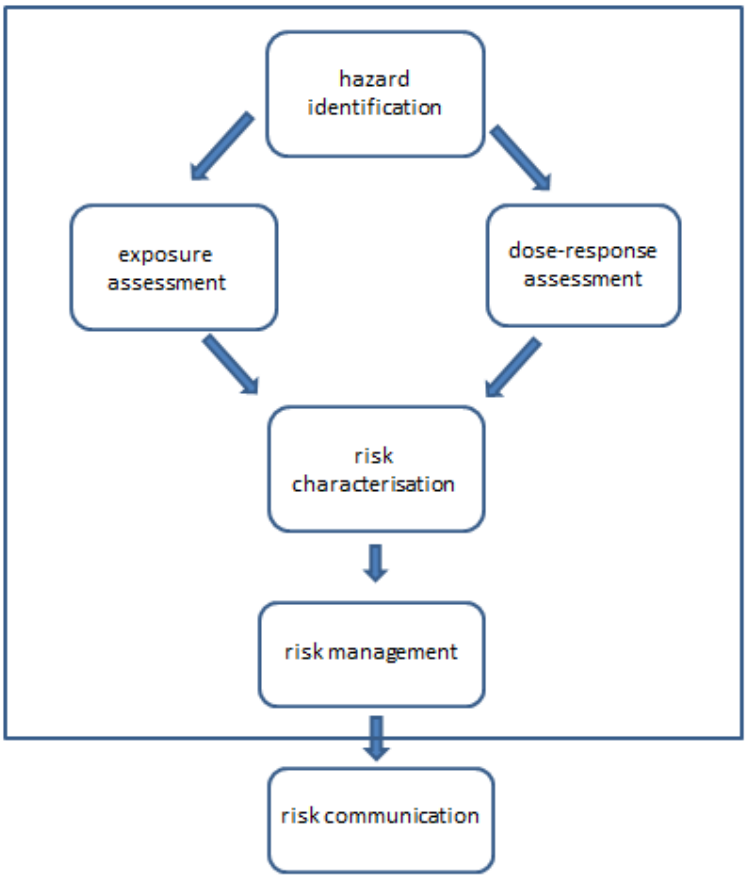

Figure 1.1. Generic risk assessment framework.

1.8 THESIS OVERVIEW

This is a thesis by publication. The body of the research comprises four papers that have been submitted to journals and are presented in the thesis as Chapters 3-6. Each paper addresses a thesis objective. In addition to an overview of the relevant literature provided in Chapter 2, each article refers to significant literature in its Introduction and Background. Similarly, the Methods for each paper are also described individually.

Chapter 2 provides a review of the literature on the major constructs and themes of the thesis: wastewater reuse and significant microbiological hazards therein, risk assessment, QMRA, BNs, the role of BNs in supporting adaptive management of environmental systems, health-based targets and an overview of previous applications of BNs in QMRA.

Chapter 3 comprises a focused exploration of the literature on the use of BNs in a context of microbial risk assessment in foodborne and waterborne pathogens,

based on the premise that BNs are emerging as an effective complementary approach that overcomes some of the acknowledged limitations of QMRA. The paper provides a comparative evaluation of the capabilities and challenges of current QMRA methods and BN models and a scoping review of recent published articles that adopt BNs for microbial risk assessments in food and water. A tabulated analysis of BN procedures and features described in the published studies is included in Table A.1 in Appendix A.

Chapter 4 was inspired by the need to conceptualise health risk modelling in the wastewater use domain. The modular system presented in this chapter was founded on the health risk modelling brief of a multidisciplinary project team tasked with validating sewage maturation ponds. The models were developed and substantiated by information from the relevant literature and further refined through consultation with expert teams from academic disciplines and regulatory authorities.

Chapter 5 presents a BN based on a quantitative risk assessment of norovirus infection associated with consumption of wastewater-irrigated lettuce. Lettuce was selected as a conservative scenario in the agricultural irrigation domain, since leafy greens are particularly susceptible to pathogen contamination during wastewater irrigation due to their large surface area, because they are often irrigated intensively and are mainly eaten raw. Norovirus was chosen to represent the enteric viruses, which are thought to be responsible for most waterborne infections in developed countries, are highly infective, often found in high concentrations in wastewaters and are resistant to treatment and persistent in the environment. The study demonstrates the utility of BNs in its efficient, visual integration of risk assessment and risk management. The ability of BNs to simulate a range of scenarios by varying exposure and risk mitigation variables and rapidly evaluate their influence on health risk endpoints is explored and demonstrated, as well as their capacity to determine obligatory conditions for optimal outcomes, minimise risk and produce predictive comparisons.

Chapter 6 describes further development and expansion of the QMRA-based BN modelling of exposure pathways embarked on in Chapter 5. In this chapter quantitative risk assessments for three reference pathogens are undertaken for scenarios of recycled water irrigation of public open space. Three contemporaneous BNs representing the three significant pathogen groups for waterborne disease –

bacteria, viruses and protozoa - are developed from the risk assessments, featuring multiple in-treatment and post-treatment risk reduction steps and the capability of modelling multiple recreation-based exposure profiles. The chapter builds on the work described in Chapter 5 by further demonstrating and describing the potential of BNs in the adaptive management of waterborne microbial health risk, through their efficient modelling capabilities.

Chapter 7 summarises and discusses the major findings of this work. Possible directions for further research emanating from the work are identified and a potential future application of the models developed in the thesis is proposed.

Chapter 2: Literature review

The research described in this thesis is founded on three broad themes: wastewater reuse, risk assessment and BNs. The review begins with an overview of the context of the case study, wastewater reuse. Under this theme the justification for and public health concerns pertaining to wastewater reuse are outlined. Next, under the risk assessment theme, the literature relating to the foundation, framework, tools, origins and types of microbial risk assessments is explored and the limitations of current approaches are examined. A discourse on traditional wastewater treatment indicators, the advent of health based targets and a brief discussion of current opinion regarding concurrent microbial exposures conclude the review of literature concerning risk assessment. The third theme begins with a brief outline of Bayesian statistical methods and taxonomy and Bayesian approaches to QMRA. BNs are defined and discussed in greater detail and lastly, the literature at the nexus of BNs and QMRA is examined, to inform the first objective of the research. This appraisal clearly establishes the novelty of the application of BNs in QMRA, for modelling human health risk associated with water recycling and leads to Chapter 3, in which a detailed exploration of existing applications of BNs in QMRA is undertaken.

As the thesis includes chapters published in peer-reviewed journals, there may be some overlap between the literature discussed in this chapter and themes that are explored in greater depth in the chapters representing the body of the research. This chapter can therefore be regarded as an overview of the concepts and precepts underpinning the research.

2.1 WASTEWATER REUSE

The total volume of water in the global hydrologic cycle is several times more than is needed to sustain the current world population, however geographic and seasonal variation results in only one third of this water being available for human use (Postel, 2000, Shiklomanov, 2000). The collective impacts on water resources of population growth, increased water requirements and regional water scarcities due to climate change have resulted in the development of water reclamation and reuse schemes (Asano et al., 2007, Postel, 2000, Shiklomanov, 2000). In Australia it is

estimated that almost 90% of rainfall is absorbed by the soil and only 12% of rainfall runs off and is collected in rivers (Radcliffe, 2004). In a highly variable climate, and with continuing population growth, Australian water authorities face increasing difficulty in providing secure water supplies. The potential for water recycling as an additional water resource has rapidly gained recognition. In 2013, an estimated 268 sewage treatment plants across Australia supplied a total of 290 GL of recycled water (BOM, 2016), representing an increase of 58% since 2009-2010 (NWC, 2014).

Recycling of waters that have previously been regarded as unusable serves a dual purpose. It can provide additional sources of water for a range of purposes, including many that are unnecessarily supplied by limited freshwater resources and it can also reduce discharge of wastewater into pristine, potable or sensitive receiving environments, such as rivers and oceans (NRMMC-EPHC-AHMC, 2006). Thus, use of treated wastewater for irrigation, cleaning or industry reduces pressure on potable water supplies and also has the potential to reduce energy and resource use through customisation of treatment requirements to provide a fit-for-purpose resource. Due to water scarcity, the practice of reclaiming and reusing wastewater is increasing in prominence worldwide, particularly in arid regions (Bitton, 2005, Drechsel et al., 2015). Recycled water can be used for a wide variety of purposes and in principle at least, the designated use is governed by the standard to which the water has been treated, although in practice, the treatment and resulting water quality are often governed by established criteria (Radcliffe, 2004). Potential uses for reclaimed wastewater include agriculture, landscape irrigation, groundwater recharge, recreation, nonpotable urban supply, potable reuse and industry (Asano and Tchobanoglous, 1991, Bitton, 2005). In Australia and the United States, agricultural and landscape irrigation are the largest uses for reclaimed water (Radcliffe, 2004, Asano et al., 2007).

In Australia to 2006, more than 270 agricultural schemes were reported to be using reclaimed wastewater and at least a further 230 schemes involving municipal application of recycled water (NRMMC-EPHC-AHMC, 2006). Possible reuse scenarios proposed in the Australian Guidelines for Water Recycling (2006) include garden irrigation, municipal irrigation, food crop irrigation (home-grown and commercial), toilet flushing, washing machine use, firefighting and dual-reticulation systems. Recommendations are published in the guidelines for these applications;

however the specified aim of the guidelines is to promote these and other uses of recycled water (NRMMC-EPHC-AHMC, 2006). In addition to safety, issues for consideration in reusing wastewater include reuse opportunities, societal acceptance, economic considerations, reliability of supply, storage during seasons when irrigation is not required, public policy and regulatory factors (Asano et al., 2007, Garcia and Pargament, 2015).

2.1.1 Health risks associated with wastewater

Treated wastewater represents a complex mixture of both microbial pathogens and chemical contaminants, with concentrations that can vary substantially depending on many parameters. In wastewater reuse, pathogens and chemicals represent two distinct categories of health hazards and their risk analysis methods are quite different. Ideally, comprehensive risk assessment of treated effluent should consider both entities, as well as the potential for interactions. Chemical contaminants of concern in wastewater include pesticides, heavy metals, halogenated compounds and other xenobiotics. Many of these chemicals are known mutagens or carcinogens and/or endocrine disruptors (Bitton, 2005). The nature and source of the influent waters are important clues in understanding the chemical characteristics of wastewater. Known and unknown inorganic and organic constituents may be present in wastewater, from the natural water supply, from stormwater incidentally or by design, added in from domestic and industrial activities or formed during treatment as a result of biotic and abiotic reactions (Asano et al., 2007). Chemical compounds are generally present in relatively small quantities in wastewater and although they are hazardous with long term, regular exposures, they are thought to present little acute risk if ingested inadvertently in minute amounts (Radcliffe, 2004). Presently there are few references in the literature to postulated interactions between chemical and microbiological contaminants in wastewater. Muñoz et al. (2010) examined chemical and microbiological contaminants in wastewater for irrigation but were unable to compute cumulative health risk estimates for the contaminated water because of lack of available disability-adjusted life year (DALY) values for many microorganisms, as well as for the organic pollutants included in the study. More recently, Genthe et al. (2013) concluded that individuals exposed to wastewater-contaminated river water had an increased probability of infection from waterborne diseases due not only to excessive microbial exposure, but also to the immune-

compromising effect of metals found in wastewaters. Varela and Manaia (2013) raise concerns about other clinically relevant issues such as the selection and spread of antibiotic-resistant or virulence genes within the indigenous microbiota in wastewater environments.

Consideration of the chemical contaminants in treated wastewater and their potential for interactions with microorganisms was considered beyond the scope of this program of research. However, the risk assessment methods applied in this project should be equally applicable to health risks related to chemicals as to microbial risk.

2.1.2 Pathogens of public health significance in wastewater

Worldwide, water- and excreta-related communicable infections include a diverse array of diseases such as Japanese encephalitis, dengue, leprosy, schistosomiasis and African sleeping sickness, in addition to more commonly recognised waterborne diseases such as diarrhoea, enterocolitis, hepatitis and cholera (Mara and Feachem, 2003). In Australia, serious illnesses such as meningitis, myocarditis, septicaemia, reactive arthritis, Guillain-Barré syndrome and haemolytic uraemic syndrome, as well as minor acute infections such as gastroenteritis, dysentery, pneumonia and skin, eye and ear infections may result from exposure to waterborne and faecal pathogens (NRMMC-EPHC-AHMC, 2006). The effects of infection from waterborne pathogens of public health significance may be mild or severe and acute, delayed or chronic. In some cases, multiple effects can result from exposure to any one microbiological hazard - for example *Campylobacter* may cause gastroenteritis, Guillain-Barré syndrome and/or reactive arthritis (WHO, 2008).

The waterborne pathogens of concern in a particular region are determined by such factors as geographical and environmental dynamics as well as the social, economic and sanitary standards of the community (Toze et al., 2012). The Queensland Water Recycling Guidelines (2005) list the following bacterial pathogens commonly found in sewage: *Salmonella* spp., *Shigella* spp., *Vibrio cholera, Clostridium* spp., *Campylobacter jejuni*, *Legionella* spp. and toxigenic strains of *Escherichia coli (E. coli*). Common viral pathogens found in sewage include enterovirus (e.g., poliovirus, coxsackievirus, echovirus and hepatitis A), reovirus, rotavirus, adenovirus and norovirus. The two most common parasitic protozoa found in sewage are *Giardia* spp. and *Cryptosporidium* spp. Helminth

parasites include tapeworms or cestodes (e.g., *Taenia saginata* and *T. solium*); roundworms or nematodes (e.g., *Ascaris lumbricoides*) including hookworms (e.g., *Ancylostoma* sp. and *Necator* sp.), whipworms (e.g., *Trichuris trichiura*), pinworms (e.g., *Enterobius vermicularis*) and flukes or trematodes (e.g., *Schistosoma mansoni*). Helminth eggs are not expected to be found in appropriately treated recycled water in Australia as they are typically removed by conventional sewage treatment (EPA QLD, 2005).

2.2 RISK ASSESSMENT

Risk is the product of likelihood of a hazardous event and magnitude of the consequence and is therefore not a directly measurable attribute (Pollino and Hart, 2008). However in the context of human health, risk is usually understood to mean an objective measure of probability of loss in terms of injury, illness, or death, following a defined event. An equally important factor in the evaluation of risk is the level of risk that is generally agreed to be acceptable or tolerable, a subjectively determined measure with potential for revision. Acceptable risk can be determined by such factors as arbitrarily defined probability, level of risk or disease which is already tolerated, or opinion of public health professionals, general public or politicians (Hunter and Fewtrell, 2001). Risk assessment is therefore a quantitative or qualitative determination of likelihood of adverse consequences resulting from exposure to hazards, with regard to tolerable risk level. Quantitative risk assessment entails consideration of the magnitude of the adverse outcome and the probability of its occurrence. The science of quantitative risk assessment is increasing in complexity; improved research methods are producing a profusion of data, leading to increasingly complex questions, such as risks in vulnerable subpopulations and how to assess multiple risks (NRC, 2009). Risk assessment in its simplest form consists of some or all of four fundamental steps:

1) hazard identification, 2) exposure assessment, 3) dose-response assessment and 4) risk characterisation (NRC, 1983). Under the original framework proposed by the United States' National Research Council in its 'Red Book', the question underpinning the risk assessment process was

"What is the probability and what are the consequences of an adverse health or ecologic effect as a result of the exposure?"

(NRC, 1983). The original framework, often referred to as the chemical risk paradigm, has since been revisited (NRC, 2009). This revised risk assessment framework begins 'with the end in mind' (Covey, 1990). The initial step now comprises enhanced problem formulation and scoping, in which risk management options are identified upfront, along with the types of technical analyses needed to evaluate the options. The updated framework asks

"What options are there to reduce the hazards or exposures that have been identified and how can risk assessment be used to evaluate the various options?"

(NRC, 2009). The original framework has been extended to include risk management and risk communication steps and has also been broadened to account for the dynamic and epidemiologic features of diseases resulting from microbial infection (Fewtrell et al., 2001).

An important aspect of quantitative risk assessment is that whole systems are envisioned, with each possible adverse event followed through to its consequences (Parsons et al., 2005). Nevertheless a quantitative risk assessment model can be built to represent the whole or part of a process or system; for example in the case of the classic 'farm-to-fork' food processing chain, the complexity and magnitude of the processes, variables and data required to populate a model representing the entire sequence may not be possible. Modelling part of the system will nonetheless produce valuable new information about the process (Parsons et al., 2005). A disadvantage of modelling selected components of a system however, is that validation of the results of the partial model against information from other parts of the system cannot be undertaken, leading to a loss of understanding of the interactions between the different modules (Albert et al., 2008).

2.2.1 Uncertainty and variability in risk assessment

Precise measurement of the two quantities comprising risk estimates – degree of potential loss and probability of the occurrence of loss - can be difficult to achieve and the chance of error in measuring these two concepts is large. Expressing uncertainty and variability is therefore not only fundamental to the risk assessment process (NRC, 2009), but the crucial point of risk assessment (NRC, 2009, Vose, 2000). Furthermore, the precision and consequent usefulness of a numerical risk

assessment rests on its ability to indicate, separate and evaluate the uncertainty and variability of the estimate (Lammerding, 1997, Vose, 2000).

Uncertainty in risk assessment refers to the degree of precision with which a quantity is measured and is introduced by factors such as the quality, quantity and relevance of data, as well as the reliability and relevance of models and assumptions. Sources of uncertainty in QMRAs include modelling pathogen densities, selection of parameter distributions, information on dose-response relationships and the choice of model (USEPA, 2010). In environmental settings, uncertainty is often the result of imperfect knowledge of the relationships between variables or processes within a system and poor understanding of the innate variability associated with environmental processes (Pollino and Hart, 2008). Uncertainty can be reduced by model refinement, including altering model parameters or by the collection of more data. Quantitative uncertainty analysis endeavours to analyse and describe the degree to which a calculated value may differ from the true value, sometimes using probability distributions (USEPA-USDA/FSIS, 2012).

Variability refers to naturally-occurring temporal, spatial, or between-individual differences in parameter values (Cullen and Frey, 1999) and can be observed for instance, in temporal and spatial differences in pathogen concentrations or variations in intake volumes between individuals (USEPA, 2010, Haas et al., 1999). Variability, as a feature of natural systems, can be examined and described, but cannot be reduced (NRC, 2009). Separation of variability and uncertainty therefore enables the risk assessor to determine whether collecting more data will produce a more accurate risk assessment. The separation and characterisation of the uncertainty and variability of model parameters is now widely recommended in risk assessment (FAO/WHO, 2003, Nauta, 2000, Vose, 2000).

2.2.2 Probabilistic risk assessment tools

Risk modelling is now an essential element in the risk management process of modern water recycling schemes and advocated in many national and international recycled water guidelines (Hamilton et al., 2007). The use of probabilistic risk assessment tools and techniques can help in the analysis and portrayal of uncertainty, variability and the impact of data gaps in risk analysis, by providing estimates of the range and likelihood of a hazard, exposure or risk. Probabilistic risk assessment employs methods to address uncertainty in scenarios, uncertainty in models and both

variability and uncertainty in the inputs and outputs of models. Such tools include Bayesian inference, sensitivity analysis, Monte Carlo (MC) simulation and expert elicitation (USEPA, 2014b). Significantly for this thesis, the United States Environmental Protection Agency (2014a) predicted that hierarchical modelling would play an important role in the future of microbial risk assessment and that Bayesian techniques may be used in the future. The use of Bayesian techniques has been anticipated to develop dose-response models in the absence of human data, for parameter estimation in the case of inadequate data sets, for data sets exhibiting wide variability, or for exposure assessment (USEPA, 2014a).

2.2.3 Quantitative microbial risk assessment

Quantitative microbial risk assessment (QMRA) is a framework for assessing public health risks from pathogenic organisms in food and water. The term describes a structured approach that brings information and data together with mathematical models to examine the exposure and spread of microbial agents and to characterise the nature of the adverse outcomes (CAMRA, 2013a, Havelaar, 2012, USEPA-USDA/FSIS, 2012). A quantitative risk assessment framework is informative in the consideration of pathogenic risk because of the many permutations of potential pathogens, sources of contamination and environmental influences which renders examination of all scenarios via epidemiological methods impossible. QMRA can be used to predict relative risks for alternate scenarios and to evaluate efficacy of management options (Soller, 2008).

The QMRA process is based upon the classic ‘epidemiological triangle’ comprising the pathogen, the human host and the environment in which exposure takes place (Cooper and Olivieri, 1998). Following identification of the ‘hazard’ or pathogen of interest (Figure 1.1), the dose response and exposure assessment steps of QMRA can be envisioned in the epidemiological triangle as the pathogen-host interaction and the environment-host and environment-pathogen interactions, respectively (Figure 2.1). Current practice in QMRA incorporates these entities into a mathematically-based method that estimates the risk associated with exposure to reference pathogens in the key pathogen groups of interest in waterborne disease: bacteria, viruses, protozoa and helminths. QMRA models are used to generate knowledge about the propagation of microbiological hazards along the risk pathway from source to exposure in complex real-world scenarios. The primary purpose of

these models is to generate insight into the interdependence of variables (input and output) and to quantify the effect of mitigation alternatives (Greiner et al., 2013). Although initially QMRA did not take into account the varying levels of severity of outcome associated with different pathogens, the health outcome metric disability-adjusted life years (DALYs) is now commonly used in conjunction with QMRA risk estimates by agencies such as the World Health Organisation (WHO), to increase the utility of the method (Cook et al., 2011).

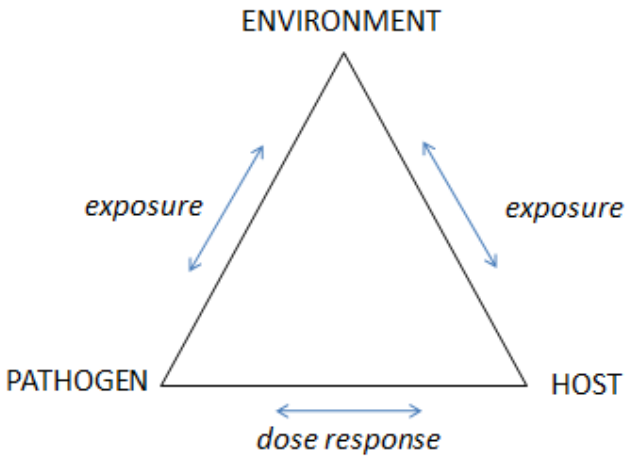

Figure 2.1. Epidemiological triangle with dose-response and exposure steps of QMRA superimposed

QMRA is is used worldwide by government agencies to protect public health from harmful exposure to waterborne and foodborne pathogenic organisms and to inform policy and decision making regarding these exposures (WHO, 2008). The QMRA framework has been applied to evaluate and manage pathogen risks for various risk pathways in drinking, recycled and recreational waters, food and the use of biosolids (Soller, 2012). QMRA is useful where no epidemiological data exist, until new epidemiological studies are developed and/or where epidemiological studies may not be practical or appropriate (Soller, 2008). Haas (2002) envisaged in 2002 that the use of QMRA would eventually replace the use of indicator-based approaches to regulation of water quality and in Australia this has occurred for recycling water guidelines, although drinking water guidelines are still indicator-based (Bichai and Smeets, 2013).

The development of QMRA

Epidemiological studies were the first documented efforts at establishing links between waterborne pathogens and disease (NRC, 2004) and historically, epidemiology has been the major scientific discipline to study the transmission of infectious disease through water (NRC, 1999). However in order to capture the key components of risk - chance, hazard, exposure and consequences, it is necessary to have replicable and scientifically justifiable methods (CAMRA, 2013b). The spread and control of measles (Hamer, 1906) and malaria (Ross, 1911) were the first infectious diseases to be mathematically modelled. Quantitative methods to characterise the human health risks associated with exposure to pathogens were published in the 1970s (Dudley et al., 1976, Fuhs, 1975). The discipline of QMRA has grown exponentially since that time, differentiating to the study of waterborne pathogens and foodborne risk assessment (Soller, 2008). QMRA explicitly and quantitatively characterises the nature and source of the pathogen, its fate, transport and kinetics, the distribution of the organisms in the medium and variables in the exposure scenario. These components are implicit in epidemiological studies, which describe demographic, spatial and cause and effect patterns of disease (Whelan et al., 2014).

The process of empirical evaluation of health risk from exposure to pathogenic microorganisms is, in general, based upon the fundamental four-step chemical risk assessment process described earlier (NRC, 1983), with some important differences. The unique features of a dynamic infectious disease process which are not accounted for in the original chemical risk framework include microbial growth and death, host immunity and susceptibility, the potential for secondary transmission, a range of possible health endpoints including delayed and/or chronic health effects, genetic diversity of microbial strains and their responses to interventions, detection method sensitivity and multiple and sequential routes of exposure. Population, community and ecosystem dynamics and heterogeneous spatial and temporal distribution in the environment are other features of risk analysis involving microbial processes and systems which require consideration (USEPA-USDA/FSIS, 2012)

The QMRA process

Six tasks of QMRA reflect the contemporary chemical risk assessment framework proposed by the United States' National Research Council (2009): hazard

assessment, dose-response assessment, exposure assessment, risk characterisation, risk management and risk communication. An alternative protocol is that published by the International Life Science Institute (2000), which more fully integrates the policymaking and technical tasks of QMRA. In the problem formulation phase, the context of the risk assessment is established, in which pathogens to be modelled and conditions to be investigated are determined. The outcome of interest is defined and the factors known to affect this outcome are identified (Alston et al., 2012). The goals and regulatory and policy contexts are established and a conceptual model is developed. Conceptualisation of the risk pathway, beginning at the source of the hazard and ending at the significant undesirable consequences, has been described as the 'backbone' of every microbial risk model (Smid et al., 2010). In the analysis phase, exposure scenarios are identified, documented and evaluated and health effects evidence, such as dose-response and epidemiological data, is gathered. The appropriate tools or methods needed to characterise risk are identified. In the risk characterisation phase, information and data are integrated and the results placed in context for the communication of risk (ILSI, 2000, Soller, 2012).

The minimum amount of data required for a feasible QMRA to calculate risk is the infective dose, concentration in the medium and the percentage removed during processing, treatment or amelioration (Soller et al., 2004, Thoeye et al., 2003). Although QMRA does not, by definition, preclude variables, pathways and population dynamics affecting pathogen densities preceding exposure, it generally focuses on factors influencing health risk at the time of exposure (Whelan et al., 2014). Researchers undertaking QMRA also commonly focus on a limited number of components and highly specific questions regarding the infection pathway (Albert et al., 2008, Soller et al., 2004). Examples of focused research questions include investigation of recreational waters to determine levels of pathogens present, investigation of the relationship between the dose ingested and the rate of illness in human volunteers for a specific pathogen strain, comparison of predicted risks with observed levels of disease and determining the impact of disease on sensitive subpopulations (Soller et al., 2004).

Traditional versus health-based targets for microbial water quality

The 'germ theory' of disease and subsequent recognition of the link between faecal pollution and waterborne disease transmission in the latter half of the 19th

century marked the beginning of modern sanitation and water treatment methods. Evaluation of microbial water safety began with the total coliform group as the routine microbial indicator in the first half of the 20th century. Although the frequency of waterborne disease outbreaks attributable to bacterial pathogens decreased, during the 1970s and 1980s there was an apparent rise in the proportion of waterborne outbreaks attributable to enteric viruses and Giardia lamblia (Sinclair et al., 2015). As detection methods for viral and protozoal pathogens were developed and improved it began to be apparent that viruses could be detected in treated water supplies that complied with existing water quality standards, challenging the assumption that the absence of coliforms was a reliable indicator of microbial water safety (Sinclair et al., 2015). Reference pathogens are now used, in conjunction with traditional indicators and surrogates, to monitor water and wastewater quality (NRMMC-EPHC-AHMC, 2006), although when used alone, these do not provide direct indication of health outcomes for exposed individuals or populations.

Worldwide, two recognised QMRA-based health outcome targets are in common use with reference to evaluating microbial water quality. The USEPA has nominated a rate of one waterborne pathogen infection per 10,000 people per year to guide calculation of water treatment requirements (USEPA, 1989) and WHO (2008) has adopted a pathogen-specific acceptable burden of disease measure of one DALY per million persons per year. The suitability of both benchmarks has been the subject of ongoing debate (Haas, 1996, Mara, 2011, Sinclair et al., 2015). The former benchmark was determined using the waterborne disease burden already tolerated in the United States, divided by the United States' population in 1987 (Lechevallier and Buckley, 2007). This, however, may not be a suitable yardstick for other countries. The burden of disease benchmark also needs to be used with full understanding of the assumptions inherent in and data required for calculation (Sinclair et al., 2015). For example, the proportion of the population susceptible to developing disease symptoms following infection is calculated from country-specific epidemiological data, which does not exist for all countries, or indeed for all pathogens. It is for these reasons that both benchmarks are called into question and arguments for revising current tolerable risk levels have been made in the literature. Furthermore, differences in the prevalence and concentration of pathogens in source waters, environmental conditions and the susceptibility of different populations often mean

that the results of one QMRA study cannot be compared with others (De Keuckelaere et al., 2015).

Sinclair et al. (2015) provide a contextual discussion for decision making regarding the selection of health-based targets, drawing attention to 'the relative and judgemental nature of defining tolerable risk'. In this paper the origins of commonly used health-based targets and their alternatives are explored and critiqued. The authors maintain that the choice of an infection risk or a DALY health target for water and wastewater-related exposures is at the discretion of the user and suggest factors to be considered in choosing a risk benchmark might include the relevance of the target to the particular setting, with due consideration given to the origins of the numeric values suggested as targets, for full understanding of inherent limitations and assumptions.

The point is further made in this discussion that chosen targets are not immutable, but require review as knowledge of disease impacts increases. For example, there is growing evidence of chronic or longer term sequelae occurring as a result of certain enteric infections such as *Campylobacter*. Recent research indicates that in addition to the known sequelae of Guillain–Barré syndrome and reactive arthritis, inflammatory bowel disease or irritable bowel syndrome may also be triggered by *Campylobacter* in some patients. Sinclair et al. (2015) note that if irritable bowel syndrome were included in DALY calculations, the health burden for an average case would increase by more than four times the current estimate and if the current target of 10^{-6} DALYs per person per year is retained, pathogen removal requirements for water exposures would increase accordingly, although these trends may be offset to some extent by improved medical management of enteric infections and their sequelae.

Water regulatory authorities have traditionally set prescriptive criteria for the microbiological quality of recycled water and performance of process trains for defined uses. In most cases such criteria are fixed microbial standards based on indicator organisms (e.g., *E. coli,* <100 CFU/100 mL) and in some cases pathogens. However the 2006 WHO guidelines for use of wastewater use in agriculture replaced the traditional approach to water quality standards with a risk assessment/risk management perspective, including the use of DALYs and a more flexible position to facilitate safe water recycling. In addition to distinct performance targets,

combinations of individual treatment components and onsite risk reduction strategies can be designed to achieve the performance target for a water recycling scheme (Huxedurp et al., 2014). A comparable approach has been used in the development of the Australian guidelines for water recycling (NRMMC-EPHC-AHMC, 2006).

QMRA endpoints, which are usually estimates of risk for infection, illness or mortality, characteristically take the form 'number of infections per person per time period' and appear to be conclusive risk predictions. However, in view of the number of assumptions and uncertainty in the modelling process, Roser et al. (2006) caution that despite QMRA risk estimates being based on scientific data, they should be viewed by decision makers as approximations of potential health outcomes or ordinal gauges of safety, rather than a precise projection. With this in view, use of BNs based on the QMRA modelling process is an expedient approach to evaluating risks and benefits of alternate control scenarios or treatment processes, or comparing the severity of different events based on their relative risks. Nonetheless, risk estimates derived from QMRA process models are frequently compared with established benchmarks in the literature.

In addition to the debate about the suitability of the two major health outcome targets described above, two other issues pertaining to health targets have been mentioned in the literature (Barker, 2014b). If different model endpoints, such as annual probability of infection, annual probability of illness and disease burden, are reported together, contradictory conclusions from the same model are possible, such as one endpoint suggesting that the modelled scenario is safe, while another endpoint signifying that tolerable risk has been exceeded. For example, in a study of greywater irrigation of home-produced lettuce (Barker et al., 2013a), although median risk estimates met the 10^{-6} DALY target, use of the 10^{-4} annual probability of infection health target resulted in all greywater irrigation scenarios being judged as unsafe (Barker, 2014b). The second issue concerns the representation of uncertainty in health targets, such that although both the modelled risk or burden of disease and the 'yardstick' or tolerable risk against which it is evaluated both contain significant uncertainty, it is customary to characterise uncertainty in the former quantity, but the latter is denoted as a point estimate (Barker, 2014b).

Types of QMRA

The type of QMRA to be undertaken is determined by the scope of the problem and the goal or required outcome. A screening QMRA is usually a conservative estimate of possible risk based on available data, which is generally simple and able to be achieved rapidly. A risk ranking QMRA may evaluate risk among several hazards, such as a single pathogen evaluated in multiple wastewater reuse scenarios, a single water source containing multiple pathogens or multiple pathogens and multiple source waters, treatment types or reuse options. A product pathway assessment identifies the key factors affecting exposure including the potential impact of mitigation strategies on predicted risk. A risk-risk analysis considers the trade-off of one risk for another and a geographic risk assessment examines the factors which either limit or enhance risk in a given region (USEPA-USDA/FSIS, 2012).

A QMRA may be static or dynamic. In a static risk assessment, an assumption is made that the number of individuals that are susceptible to infection does not vary with time and risk is characterised at an individual level (Eisenberg et al., 2002). In a dynamic risk assessment, the number of individuals assumed to be susceptible to infection varies with time and risk is analysed at the population level. In the latter, issues such as person-to-person transmission and immunity can be taken into account (Anderson and May, 1991, Hethcote, 1976, Hethcote, 2000). Dynamic microbial risk assessment models may be further defined as deterministic or stochastic. Deterministic dynamic models are comprised of a set of differential equations with defined parameters and starting conditions that produce point estimates (Eisenberg et al., 2005, Soller et al., 2003). One shortcoming of the deterministic approach is that selection of worst case ‘safe’ estimates for each variable results in propagation of conservatism throughout the system, resulting in a worst possible combination of events which overestimates exposure and is highly improbable. Moreover, the likelihood of the estimated exposure actually occurring is unknown (FAO/WHO, 2008).

Stochastic or probabilistic risk assessment models incorporate probability distributions and undergo iterative evaluations to reach convergence (Koopman et al., 2002). Stochastic models represent all the information available for each variable, described as a probability distribution of possible values. The outcome of a stochastic

model is a statistical distribution that describes the range of possible outcomes and the likelihood that they will occur (FAO/WHO, 2008). Although dynamic models are capable of approximating biological 'realism', they are analytically complex, resulting in increased computational demands and variability due to the uncertainties associated with model specification (USEPA, 2014a). Conversely, a simpler model form involves implicit or explicit assumptions that may or may not be realistic or appropriate for a particular situation. (Soller and Eisenberg, 2008) determined that a simpler static model provides satisfactory risk estimates under low risk conditions (e.g., 1 in 1,000 or 1 in 10,000), as defined by a combination of exposure levels and infectivity of the pathogen.

QMRA limitations

Microbial risk assessments are largely comprised of numerical simulation studies and therefore need to be founded on empirical data to be widely accepted (Soller, 2012). However the apparently straightforward process of counting microscopic pathogens is subject to numerous influences, some of which are systematic and may be quantifiable and some that are random (Crainiceanu et al., 2003). Another limitation of QMRA studies is that they are pathogen-specific, as opposed to the broader scope of epidemiologic studies.

A quantitative assessment of the overall health risk of exposure to waterborne pathogens, taking into account the different exposure routes, is not currently feasible, as the data relating to exposure by inhalation or dermal contact for example, are scarce or not available (O'Toole, 2011). Furthermore, in a scenario with different exposures occurring simultaneously, it is difficult or even impossible to determine which exposure route is responsible for which proportion of the risk. Quantitative health risk assessments are thus presently based on the most risky exposure route, which is usually direct ingestion in the case of recycled water, as it delivers the largest dose of pathogens (EPA QLD, 2005). This assumption may vary depending on the reuse scenario - for example, in the case of spray irrigation the most risk-laden route may be inhalation or contact with skin of aerosolised wastewater (Thoeye et al., 2003). Due to the low numbers of microorganisms (tens or hundreds) involved in exposure to water, there may be large differences with respect to the actual number of organisms ingested between individuals (Haas, 2002). Owing to the significant variability and uncertainty associated with assessing absolute risk from exposure to

pathogenic microorganisms, QMRA is considered easiest to conduct and clearest to interpret when comparing the relative risk of two or more scenarios (Soller et al., 2004). Apart from the difficulty in characterising exposure, other problems inherent in the QMRA process include subjectivity in model and parameter selection, the significant effect of differential susceptibility in the host (together with little quantitative information) and the scarcity of dose-response data (Soller, 2008). The limitations of existing dose response information include the use of healthy volunteers and/or attenuated or organisms in experiments, omission of low doses, the use of culture-based methods of enumeration and the definition of response varying from faecal excretion, to antibody response and sometimes symptomatic illness (Rose et al., 2008).

These limitations in available data and models are acknowledged by the World Health Organisation (2008). Vose (2000) notes that although QMRA has been taken up on a global basis, '… the popularity of risk analysis does not seem to have been matched with corresponding improvement in the understanding of its techniques'. Medema and Ashbolt (2006) assert that although QMRA is a suitable tool to assess the potential health risk of water-related systems, it is less appropriate in the assessment of actual health risk of drinking water consumers. In all, the QMRA concept and framework is widely recognised, but there is scope for current approaches to be enhanced by additional techniques (Havelaar, 2012, WHO, 2008).

Concurrent microbial exposures

As stated earlier, QMRA is a pathogen-specific modelling approach because the specificity of the dose-response relationship is based on experiments using single pathogen types. Wastewater however, whether treated or untreated, contains an assortment of microorganisms, which enter the host collectively. Population dynamics such as competition, predation and other relationships such as symbiosis are theoretically possible in mixed flora, but presently there is very little conclusive information available regarding such microbial interactions and the immune response of the host in concurrent microbial exposures (Gerba, 2015, Haas, 2015, Rose, 2015).

A treatise on pathogenic exposure and immune response (Brownlee, 2007) extrapolated multiple pathogen exposure-response events from a single pathogen exposure-response model. The results indicated the efficiency of the immune system would be reduced in such a situation, as resources are allocated proportionally to the

relative virulence of each pathogen type, with the implication that infection risk would be increased with multiple concurrent pathogen exposures. This hypothesis is in conflict with a 2009 study (Pujol et al., 2009), characterising infectivity as a function of nonspecific pathogen dose. The model in this study revealed an inverse relationship between time for subsequent exposures and infection risk, suggesting exposure to one pathogen potentiates the immune system, to keep concurrently or subsequently arriving particles from initiating an infection. This study concluded that single-dose dose response experiments may overestimate infection risks in the real world. The study focused on the effect on dose-response of multiple exposures accumulating over time, as opposed to simultaneous or concurrent exposures. At the very least, the study challenges the notion of independent risk events for each pathogen exposure, also suggesting dose-response for concurrent exposures may be overestimated if risks are simply aggregated.

For the present, risk modellers handle concurrent risk estimates from separate pathogen classes by adding them together, with the implicit assumption that the immune responses are independent events in probabilistic terms. A similar procedure may not be valid for burden of disease estimates (Rose, 2015, Petterson, 2016), due to the variable quality and availability of data from different sources required to calculate DALYs and other summary measures of population health.

In subsequent paragraphs, BNs are discussed in detail, including their advantages and disadvantages, their use in adaptive management and existing applications of BNs in QMRA.

2.3 BAYESIAN STATISTICAL METHODS

The Bayesian paradigm in statistics has now become so prevalent as to be described as ‘perhaps the dominant paradigm for doing statistics’ (Alston et al., 2012). Briefly, Bayesian inference estimates the probability of a hypothesis being true given the available data, whereas frequentist inference estimates the probability of a set of data occurring given a particular hypothesis (Ellison, 2004). Bayesian statistics is an established methodology and is being applied in a broad range of research contexts, from health, the environment and ecology to finance, engineering, philosophy and the law (Alston et al., 2012). The common underlying thread for

these applications is the elementary identity of Bayes' theorem and the probabilistic expression of uncertainty about unknown parameters (Cowles et al., 2009).

2.3.1 Bayesian 'taxonomy'

The term 'Bayesian computational methods' encompasses a wide range of algorithms used in Bayesian analysis, including the elemental Markov Chain Monte Carlo algorithm and the Gibbs and Metropolis-Hastings algorithms. 'Bayesian methodology' generally implies the use of prior distributions in data analysis. Together Bayesian methodology and Bayesian computation methods form the basis of 'Bayesian modelling', which alludes to a diverse array of models including normal linear and hierarchical models, Bayesian classification and regression trees, BNs and Bayesian meta-analysis (Alston et al., 2012).

'Bayesian inference' or 'Bayesian analysis' (terms which seem to be used interchangeably) begins with a prior distribution which could be based on non-Bayesian observations, or an assessment of the relative likelihoods of parameters; data is then collected to obtain the observed distribution and the likelihood of this distribution is calculated as a function of parameter values. The likelihood function is then multiplied by the prior distribution and normalised over all possible values resulting in the posterior distribution, on which all Bayesian inference is based (Cowles et al., 2009, Pearl, 2000, Weisstein, 2013). In informal terms, Bayesian inference can be regarded as a mathematical description of the learning process, in which one starts with an opinion, which is modified when presented with evidence (Vose, 2000). Bayesian inference is one technique, in addition to classical frequentist techniques, for determining the distributions of uncertainty for the parameters of a variability model, which is the primary objective in quantitative risk analysis (Vose, 2000).

2.3.2 Bayesian approaches used in QMRA

Researchers have used Bayesian statistical methods to analyse dose response data in microbial risk assessments of both foodborne and waterborne pathogens, arriving at better risk estimates due to the ability of Bayesian methods to separate uncertainty and variability in the dose response parameters (Delignette-Muller and Cornu, 2008, Messner et al., 2001, Schmidt et al., 2013b). In a similar application Duffy et al. (2006) used Bayesian analysis to reduce the uncertainty around the

predicted risk estimate in the exposure assessment model in a QMRA of *E. coli* in a food production chain Other researchers have used Bayesian modelling to characterise and understand the spatial and temporal variability in waterborne pathogen concentration estimates with associated uncertainty from presence/absence observations (Petterson et al., 2009), or when the pathogen counts displayed sizeable variation in recovery rates and included many small counts, zero counts and missing data (Crainiceanu et al., 2003, Schmidt et al., 2013a).

2.3.3 Bayesian networks

BNs are probabilistic graphical models in which the variables, represented as nodes, are connected by a directed arc or arrow, implying causality (Jensen and Nielsen, 2007, Pearl, 1988, Pearl, 2000). Each variable has a number of user-defined states that can be qualitative or discrete (e.g., 'True/False', 'High/Low', '>5/≤ 5'). Node states have conditional probabilities assigned to them, derived from empirical data, models, simulations, or from expert opinion. Software supporting BNs comprise a suite of complex algorithms and node states and their conditional probabilities are displayed in a table underlying each node (Figure 2.2). The network of connected variables forms a directed acyclic graph (DAG), in which no feedback or loops can occur (Fenton and Neil, 2013).

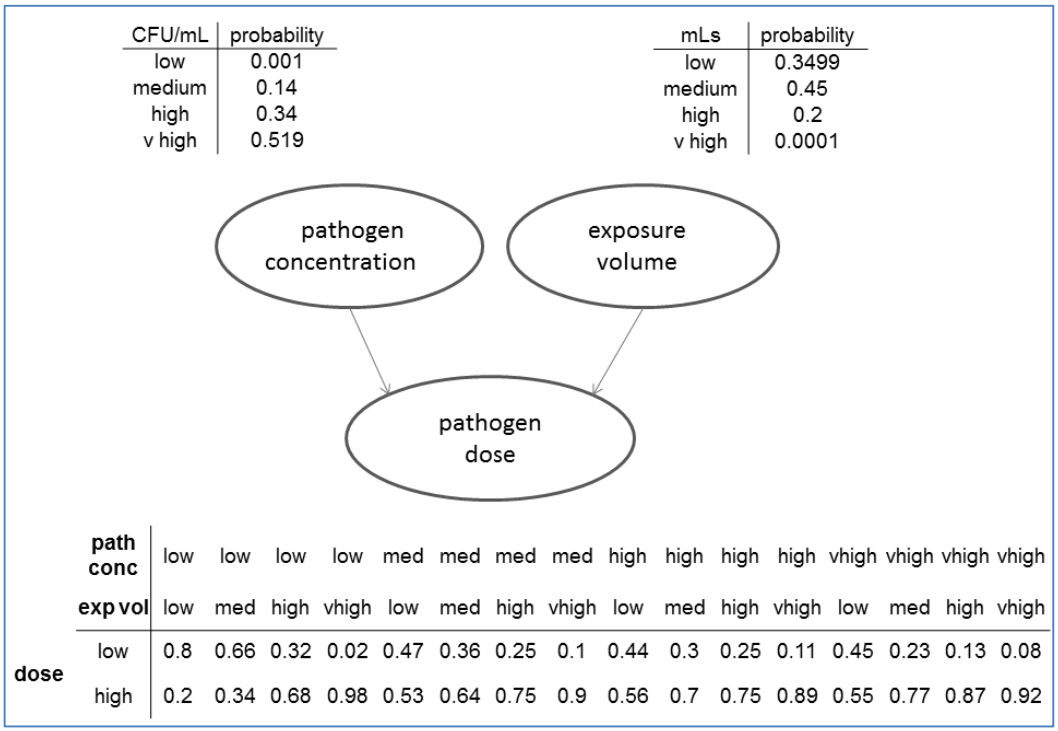

Figure 2.2. Directed Acyclic Graph (DAG) and attendant conditional probability tables.

As indicated previously, the conditional probabilities in BNs can be generated in a number of ways, including induction from databases (Korb and Nicholson, 2011), either manually or through learning algorithms in the supporting software. Conditional probabilities can also be induced from meta-analyses, models, simulation, or where data do not exist, from the structured, rigorous elicitation of expert opinion (Renooij, 2001, O'Hagan et al., 2006, Low Choy et al., 2009). Since the number of conditional probabilities in a single node table depends on the number of its states as well as the number of states in all of its 'parent' or causal nodes, if nodes have many parents and/or parent nodes have a large number of states, conditional probability tables can become very large (as the conditional probability table underneath the *Pathogen dose* node in Figure 2.2 illustrates). In environmental domains the number of conditional probabilities often exceeds the data available. BNs however, can integrate both existing data and expert opinion where data are lacking, for parameterisation and evaluation (Pollino et al., 2007). Pollino and Hart (2005) argue strongly nonetheless that use of expert opinion to parameterise environmental variables that are inherently quantitative or defining variables qualitatively when quantitative data exist, prevents objectivity in risk assessments. Expert opinion may be subject to cognitive and knowledge-based bias (Pollino and

Henderson, 2010), or knowledge of experts may be incomplete (Pollino et al., 2007). Pollino et al. (2007) also caution that rigorous elicitation of conditional probabilities in complex variables from experts can be challenging and time-consuming. Pollino and Hart (2005) further stress that historical data should be used where possible for quantifying variables, in conjunction with knowledge sources and a weighting system and importantly, iterative network updating.

BN models are thus able to characterise and quantify a complex outcome, as well as the variable interactions associated with the outcome, in cases where empirical data are sparse, missing or inaccurate (Donald et al., 2009). BNs are able to reveal variables that are major drivers for an outcome, or conversely, the sensitivity of the outcome to variables in the network (Coupé et al., 2000). A probabilistic framework such as that provided by a BN offers a systematic, holistic approach to the assessment of risk, enables integration of both quantitative and expert knowledge, provides a graphical description of the structure of the decision problem and facilitates investigation of causality (Pollino and Hart, 2005). The use of BN models does not necessarily imply a commitment to Bayesian statistics, although the use of Bayesian statistics in conjunction with a BN results in an efficient approach to fitting data. Risk analysts may follow frequentists' methods to estimate the parameters of the BN (Ben-Gal, 2007). The majority of BNs are static, but the requirement to incorporate time as a variable can be met with object-oriented and dynamic BNs (Johnson et al., 2010, Johnson and Mengersen, 2011).

The use of BNs is increasing exponentially across a wide range of application domains, such as medical diagnosis, safety assessment, the law, forensics and more (Korb and Nicholson, 2011). Technical advancements and software development mean there is now no need for end users to understand the underlying BN algorithms in order to build and use BN models, which has possibly assisted the recent rapid uptake of the approach (Fenton and Neil, 2013). Aguilera et al. (2011) reviewed 128 articles on BNs used in environmental modelling over a 21 year period ending in 2010, analysing the applications in terms of study aims and the presence of aspects such as model learning and validation. The analysis suggested that most of the research effort on BNs to date has focused on theoretical and methodological development and software implementation issues and that use of BNs in the

environmental sciences is still largely unexploited (Aguilera et al., 2011, Johnson and Mengersen, 2011).

2.3.4 Bayesian networks in adaptive management

BNs have proven to be valuable in adaptive management of environmental systems (Chen and Pollino, 2012, Laniak et al., 2013, Pollino and Henderson, 2010). Developed in response to the difficulty of making decisions under the widespread uncertainty, insufficient knowledge and constant flux characterising environmental domains, adaptive management is a structured, iterative process of learning from the outcomes of management actions, thereby progressively increasing knowledge of the system being managed, reducing uncertainty and improving the robustness of decision making (Holling, 1978, Walters, 1986).

Adaptive management leads to a shift in focus from traditional management approaches such as prescriptive regulations to increased transparency, stakeholder-driven decision making and the acknowledgement and further, embracing of uncertainty to improve understanding of a system (Henriksen et al., 2012). Eschewment of 'black box,' or inscrutable risk analysis, wherein the risk assessment process is accessible only to scientists and modellers and embracing uncertainty is a premise underpinning this thesis. BNs support the embracing of uncertainty through explicit expression in probability distributions and visual transparency. Characteristics of traditional management approaches adapted from Henriksen et al. (2012) are compared with those of adaptive management (IOM, 2013, Rajan and Letourneau, 2012) in Table 2.1. The attributes of BNs, expounded in Chapter 4 and to a lesser extent in Chapters 5 and 6, map relevantly to the features of the adaptive management approach (McCann et al., 2006).

Table 2.1

Comparison of characteristics of traditional and adaptive management approaches, adapted from Henriksen et al. (2012)

Traditional management approach	Adaptive management approach
isolation of risk factors	holistic mapping and exploration of risk factors
objectivity	subjectivity
reductionism	systems thinking
technical problem solving	collective problem solving
authority driven management	stakeholder-driven decision making
uncertainty obfuscated	transparency increased
uncertainty undesirable	uncertainty embraced

2.3.5 Applications of Bayesian networks in QMRA

A QMRA may provide quantitative inputs into a BN, a BN may be used to augment a QMRA (Donald et al., 2009), or alternatively, an entire QMRA model can be formulated as a BN in a network which expresses the joint distribution of all model variables (Greiner et al., 2013), as has been undertaken in the BNs developed in this study. To gauge the range and extent to which BNs have been implemented in conjunction with QMRA, a search of the literature via ISI's Web of Knowledge was undertaken based on the terms 'Bayesian network', 'Bayesian belief network', or 'Bayesian graphical model', used in combination with the terms 'QMRA', or 'microb*'. QMRAs for foodborne or waterborne pathogens where BNs were used, with or without the use of Bayesian statistical methods, were included in this protocol for a scoping literature review, which is presented in full in Chapter 3. Eighteen papers were selected for inclusion based on their relevance to these search criteria. Of the 18 papers examined, 14 of them were published within the last five years. Twelve articles pertained to microbial risk assessment of foodstuffs and the remaining six related to waterborne microbial hazards.

The use of BNs in food-related risk assessments included estimating the probability of campylobacteriosis (Albert et al., 2008), identification of risk factors in the food processing chain contributing to contamination (Barker et al., 2002, Liu et al., 2013, Parsons et al., 2005), biotracing (Barker and Gomez-Tome, 2013, Smid

et al., 2011, Smid et al., 2012), variability and uncertainty analysis in cross-contamination ratios (Smid et al., 2013), predictive microbiology themes such as food processing chain influences on microbial growth and survival, variability in growth rate between strains and growth parameter uncertainty (Delignette-Muller et al., 2006, Pouillot et al., 2003, Rigaux et al., 2012b) and differences in concentrations between pathogen strains (Rigaux et al., 2012a).

Of the six papers concerning waterborne pathogens, two papers described the use of BNs in assessing the health risk of recycled water (Cook et al., 2011, Donald et al., 2009), two authors used BNs in the study of recreational waters (Gronewold et al., 2013, Staley et al., 2012) one used a BN to account for variability and measurement errors in drinking water microorganism enumeration data (Schmidt and Emelko, 2011) and one paper used the technique in assessing the public health risk of wet weather sewer overflows (Goulding et al., 2012). A report by Cook (2011) proposed the use of a BN in conjunction with QMRA in a public health–based risk assessment to evaluate the safety of existing and proposed recycled water use in Western Australian communities. The relative scarceness of literature on the use of BNs in QMRA and the breadth of applications is clearly indicative of the scope for the research submitted in this thesis. These papers are discussed further in the next paragraph and analysed in greater detail in Chapter 3.

Donald et al. (2009) developed a BN as a supplementary analysis to a QMRA, which described a conceptual model for health risks associated with recycled water, with a chosen health endpoint of gastroenteritis. The model was generic in the sense that it could be used to assess the risk related to any enteric pathogen in a recycled water scheme. In the first instance, the opinion of a domain expert was used to quantify the model, but more quantitative inputs could be provided by QMRA estimates. Model predictions were presented initially in terms of probabilities and then expressed as relative risks of gastroenteritis, relative to the input nodes at the safest settings (i.e., least likely to lead to gastroenteritis). In addition to demonstrating how a BN can be used to augment QMRA, this research also contributed a methodology for quantifying the uncertainty of point estimates arising from BNs by calculating credible intervals and thus adds to the available tools for assessing microbial health risk as a result of environmental exposures. Staley et al. (2012) developed a Bayesian model using concentrations of faecal indicator bacteria

(FIB), frequency of pathogen detection and physicochemical parameters such as temperature and salinity to determine factors predictive of human health risk in a freshwater lake used for recreation. In a similar recreational water setting, Gronewold et al. (2013) used a BN to explore differences between analytical methods for quantifying FIB concentrations (most probable number and colony-forming units) and between different sampling locations and times. Goulding et al. (2012) developed a BN model with QMRA inputs to identify the public health risks associated with microbiological contaminants as a result of wet weather sewer overflows discharging into an urban waterway.

2.3.6 Advantages of using Bayesian networks in QMRA

As noted in the previous section, Bayesian belief networks are being used progressively in QMRA (Barker, 2004) and the benefits and drawbacks of the approach described in the literature are presented here.

Although the graphical structure is not a unique feature of BNs (Smid et al., 2010), representation in a BN of large quantities of complex information provides an informative platform for improved communications between mathematical modellers, subject and process experts and stakeholders (Barker et al., 2002). Furthermore, a BN which has been validated enables a clear understanding of the effect of interventions on the outcomes of interest by altering the prior distributions of variables to mimic the implementation of a new strategy (Albert et al., 2008). Interventions to reduce risk can be simulated in the network by changing parameter values, thus allowing calculation of risk reduction (Smid et al., 2010). A BN is a natural framework for combining results of a risk assessment with results from epidemiological studies, as evidence from multiple variables in the network can be used to update the estimates of the parameters in the model (Smid et al., 2010). Prior knowledge, based on historical data or expert opinion, can be incorporated with other data in the analysis (Pouillot and Delignette-Muller, 2010, Smid et al., 2010). BNs by their nature allow ‘backwards reasoning’ - downstream data combined with the prior distributions can be used to update the priors, positively or negatively adjusting the initial beliefs of the expert (Greiner et al., 2013), whereas standard MC approaches cannot use data sets downstream of other data sets (Albert et al., 2008). This feature facilitates application of specialised knowledge for forward reasoning (risk prediction) or backward reasoning (risk diagnosis) (Yang et al., 2006).

A particularly important feature for QMRA is that poor quality data in a node may have little impact on a distribution function for a key variable such as pathogen concentration, that is established from high quality prior information (Barker et al., 2002). The BN responds immediately to changes in the network, such as entering new evidence, because it does not use simulation; information is propagated from any point in the network to all others by Bayesian inference (Parsons et al., 2005). Finally, a complex, multivariate statistical problem can be efficiently addressed using available data and expert knowledge, where classical statistical methods are often inefficient (Albert et al., 2008). The efficiency of a BN framework is also evident in the systematic representation of the joint probability in such a system, significantly reducing complexity (Smid et al., 2010).

2.3.7 Limitations of using Bayesian networks in QMRA

A common limitation in Bayesian analysis is the accuracy and precision of assumed prior probability distributions (Mitchell-Blackwood et al., 2012). Another issue identified in the literature is that discretisation of continuous variables into categories or states introduces errors that can accumulate, producing an inaccurate representation of the model distributions (Parsons et al., 2005, Smid et al., 2010). However, an important development in the advancement of BNs has been the development of hybrid models, in which continuous and discrete variables can coexist (Aguilera et al., 2011). Further concerns include the inability of BNs to support feedback loops due to their acyclic nature (Jensen, 2001); this issue can be overcome by the implementation of a dynamic BN, although these models can become so large as to be computationally infeasible (Smid et al., 2011). A final drawback identified in the literature that may arise when performing exact probabilistic inference across a complex domain, is that storage of a joint distribution may become memory-intensive and therefore computationally challenging. This issue can be overcome by simplifying the joint distribution, for example by using sensitivity analyses (Smid et al., 2011).

2.4 CONCLUSION

This concludes the synopsis of literature around the central themes in the thesis. As the literature reveals, water recycling is a potential solution to increasing water scarcity throughout the world; however prudent reuse requires assurance of

microbial safety, due to the pathogens of public health significance found in wastewater. Moreover, in the interests of resource conservation, microbial safety is a criterion that, ideally, varies in accordance with the indicated purpose for the recycled water. Faecal indicator organism levels or pathogen densities in treated wastewater are inadequate when used in isolation for assessment of health risks associated with exposures. There is a widespread need for comprehensive, systematic appraisal of risk pathways, such as QMRA, for evaluation of water recycling scenarios. However, existing quantitative methods for risk evaluation are constrained by data scarcity and widespread uptake of water recycling schemes is hindered by lack of information about exposure scenarios. Finally, traditional prescriptive criteria for microbiological quality of recycled water are now being augmented with a multiple barrier approach and risk assessment outcomes are now commonly expressed in terms of health based targets as well as established microbial water quality indicators. Despite BNs having a number of features that address many of these issues, the technique has not previously been used with QMRA to assess risk in a water recycling context. These key findings from the literature have informed the objectives of this research and support the ensuing studies, described in the body of the thesis. A focused review of publications describing applications of BNs in QMRA follows in Chapter 3 and further examination of relevant literature occurs in the introduction and background sections of Chapters 4, 5 and 6.

Chapter 3: Beyond QMRA: Modelling microbial health risk as a complex system using Bayesian networks

Preamble

This chapter has been written as a journal article to meet Objective 1 of the research, as stated in the Introduction:

Objective 1 - To identify and fill a gap in the peer-reviewed literature on applications of BNs in QMRA

Exploration of existing work where BNs have previously been used with QMRA was considered an essential first step in and foundation for, the development of the models to follow. This chapter is primarily my own work and the figures and tables were created by me. The article was published by Environmental International in August, 2015 and is reproduced here in its entirety. The reference for the publication associated with this chapter is:

Beaudequin, D., Harden, F., Roiko, A., Stratton, H., Lemckert, C., & Mengersen, K. (2015). Beyond QMRA: Modelling microbial health risk as a complex system using Bayesian networks. *Environment International, 80*, 8-18.

Contributor	Statement of contribution
D. Beaudequin	Conception and conduct of research, data analysis, interpretation of results, writing of manuscript, modifications to manuscript as suggested by co-authors and reviewers
F. Harden	Comments on manuscript, editing
A. Roiko	Comments on manuscript, editing
H. Stratton	Comments on manuscript
C. Lemckert	Comments on manuscript
K. Mengersen	Comments on manuscript, editing

Principal Supervisor Confirmation: I have sighted email or other correspondence for all co-authors confirming their authorship.

Name: Kerrie Mengersen Signature:  Date: 29/04/2016

Abstract

Background: Quantitative microbial risk assessment (QMRA) is the current method of choice for determining the risk to human health from exposure to microorganisms of concern. However, current approaches are often constrained by the availability of required data and may not be able to incorporate the many varied factors that influence this risk. Systems models, based on Bayesian networks (BNs), are emerging as an effective complementary approach that overcomes these limitations.

Objectives: This article aims to provide a comparative evaluation of the capabilities and challenges of current QMRA methods and BN models and a scoping review of recent published articles that adopt the latter for microbial risk assessment. Pros and cons of systems approaches in this context are distilled and discussed.

Methods: A search of the peer-reviewed literature revealed 15 articles describing BNs used in the context of QMRAs for foodborne and waterborne pathogens. These studies were analysed in terms of their application, uses and benefits in QMRA.

Discussion: The applications were notable in their diversity. BNs were used to make predictions, for scenario assessment, risk minimisation, to reduce uncertainty and to separate uncertainty and variability. Most studies focused on a segment of the exposure pathway, indicating the broad potential for the method in other QMRA steps. BNs offer a number of useful features to enhance QMRA, including transparency and the ability to deal with poor quality data and support causal reasoning.

Conclusion: The method has significant untapped potential to describe the complex relationships between microbial environmental exposures and health.

3.1 INTRODUCTION

Quantitative microbial risk assessment (QMRA) is an established framework for assessing public health risks from pathogenic organisms (Haas et al., 1999). As a relatively recent addition to the risk analysis field (Havelaar et al., 2008), the QMRA methodology is evolving and current challenges suggest scope exists to augment established methods to improve its capabilities. Bayesian networks (BNs) are emerging as an attractive way of modelling 'wicked' problems, particularly in complex environmental systems (Aguilera et al., 2011, Barton et al., 2012, Uusitalo et al., 2012). In essence, a BN is a flexible graphical model that incorporates dependencies among its variables via probabilistic relationships. The method allows the integration of a range of quantitative information, which is particularly useful in environmental domains such as QMRA, where traditional experimental and observational data are missing, inaccurate, sparse or costly (Aguilera et al., 2011). The use of BNs is increasing exponentially across a wide range of application domains (Aguilera et al., 2011, Barton et al., 2012, Korb and Nicholson, 2011). Despite this increased interest and a growing body of literature, there has not yet been a critical review and evaluation of the approach in the context of QMRA across domains, that draws together the literature and can be used to educate and guide practitioners. This paper aims to fill that gap by exploring the range of applications of BNs in the microbial risk assessment domain.

We begin with a background of current challenges in QMRA and a description of BN models and their features in more detail. We then examine published examples of the use of BNs in QMRA and discuss the applications of the method to assess and describe human health risk in different exposure domains. Finally, we discuss advantages and limitations of the approach in the context of QMRA and attention is drawn to gaps in the reporting of current research in the area.

3.2 BACKGROUND

3.2.1 Quantitative microbial risk assessment

QMRA is a structured approach which brings information and data together with mathematical models to examine the exposure and spread of microbial agents and to characterise the nature of the adverse outcomes (USEPA-USDA/FSIS, 2012). The six steps of a QMRA, illustrated in Figure 3.1, are hazard characterisation,

exposure assessment, dose-response assessment, risk characterisation, risk management and risk communication (NRC, 2009). A typical example of the output from a QMRA model is the probability of infection or illness associated with ingestion of food or water containing pathogens. This may then be used to predict the number of cases of illness caused by pathogens introduced into a food production chain, or the incidence of waterborne disease in a population of interest (Smid et al., 2011).

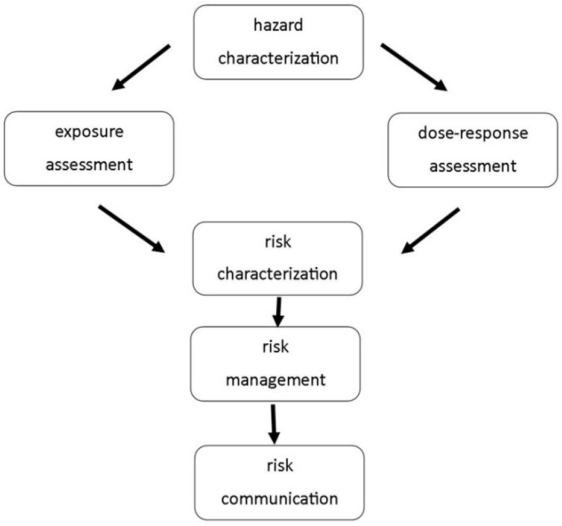

Figure 3.1. The six steps of a quantitative microbial risk assessment.

Despite their widespread use, established QMRA methods have some inherent constraints. For a feasible QMRA, the minimum amount of data needed to calculate microbiological risk is the infective dose, calculated from the pathogen concentration in the exposure medium and the quantity of medium implicated in the exposure (Soller et al., 2004, Thoeye et al., 2003). A significant limitation in QMRA however, is the lack of appropriate data to quantify established models in the dose-response step. The availability of pathogen-specific dose-response models is limited, due in part to the cost and ethical dilemmas of human feeding experiments. Where it does exist, dose response information has limitations which include the use of healthy

volunteers and/or attenuated organisms in experiments and omission of low doses from experiments (Rose et al., 2008), the significant effect of differential susceptibility in the host (Soller, 2008) and the definition of response, which may vary from faecal excretion to antibody response and sometimes symptomatic illness (Rose et al., 2008).

A second key issue arises in measuring the many facets of microbial exposure. For example, a quantitative assessment of the overall health risk of exposure to a pathogen-containing medium such as water, air or soil, taking into account all potential exposure routes, may be infeasible, as the data relating to exposure routes such as inhalation or dermal contact for example, are scarce or not available. Furthermore, in a scenario where exposures occur via multiple routes simultaneously, it is difficult or even impossible to determine which exposure route is responsible for what proportion of the risk. As another case in point, the apparently straightforward process of the detection and quantification of microscopic pathogens is subject to multifarious influences (Crainiceanu et al., 2003). For example, due to the low numbers (tens or hundreds) of microorganisms entailed in exposure from water, there may be large differences with respect to the actual number of organisms ingested between individuals, impacting model assumptions (Haas, 2002). Other sources of uncertainty and variability in enumeration data include sample representativeness, recovery efficiencies, detection limits, microbial kinetics such as resistance, die-off and growth and differentiation between strains (Petterson et al., 2006, Smid et al., 2011). Although the QMRA concept and framework is widely recognised, the limitations in available data and models are acknowledged by authorities (WHO, 2008) and there is recognition of the scope for current approaches to be enhanced and complemented by alternative techniques (Havelaar, 2012, WHO, 2008).

A third key issue is that environmental systems that can be modelled using QMRA are characterised by significant levels of uncertainty and variability, due to complex ecology, population dynamics and the multiple physicochemical and biotic influences at play. Uncertainty signifies the degree of accuracy and precision with which a quantity is measured and can be characterised and reduced by altering the model and/or collecting more data. In contrast, variability, as a feature of natural systems, can also be characterised, but cannot be reduced (NRC, 2009). The

precision and consequent usefulness of a numerical risk assessment rests on its ability to indicate, separate and evaluate the uncertainty and variability of the estimate (Lammerding, 1997, Vose, 2000). Thus, the separation and characterisation of the uncertainty and variability of model parameters, is now widely recommended in risk assessment (CAC, 1999, FAO/WHO, 2003, Vose, 2000).

There is increasing interest in viewing the microbial risk pathway as a system, such as the modular process risk modelling process (MPRM) proposed by Nauta (Nauta, 2001, Nauta et al., 2007), in which discrete processes or events in the risk pathway are represented as linked modules. A systems approach encompasses both holistic and modular views, enables the synthesis of knowledge of the parts in order to help understand the whole and makes a complex system more manageable (Auyang, 2004). A MPRM aims to model the transmission of micro-organisms along the food pathway by breaking down the pathway into consecutive modules and then modelling the basic microbial processes that take place in each module. For example, the dynamics of *Salmonella* in the 'farm-to-fork' pork slaughter chain can be described by the six basic MPRM processes of growth, inactivation, mixing, partitioning, removal and cross-contamination. At least one of these six basic processes is assigned to each key step in the chain, such as killing, scalding and dehairing, for modelling purposes (Smid et al., 2011).

BNs are one of a range of modelling tools that offer a systems perspective, with the added advantage of conditional dependence of the modules. Although other techniques such as Monte Carlo (MC) simulation of standard QMRA models share commonalities with BNs, including the expression of parameter uncertainty using distribution functions and visualisation as network graphs, a BN conveniently infers immediate changes in parameter values when new evidence is added (Greiner et al., 2013). The attraction of BNs includes the ability to address the three key issues outlined above: the scarcity of dose-response data and uncertainty in dose-response models, the difficulties with modelling exposure pathways due to complexity and lack of data and the necessity, in order to produce an informative risk estimate, to characterise and separate uncertainty and variability. In addition BNs offer a number of other features which are useful in QMRA and which are described below.

3.2.2 Bayesian networks

A BN is a form of graphical model with variables represented by nodes and connections between the variables represented by directed arcs (Jensen and Nielsen, 2007). Each node category or 'state' is assigned a probability distribution conditional on its parent nodes; these distributions can be derived from empirical data, statistical models, simulations, published papers or reports, or from expert opinion (Pollino et al., 2007). Arcs linking the nodes represent dependencies, with the strength of the causal links represented by these conditional probabilities. The directed arcs and the constraint that the arcs cannot form cycles or feedback loops within the model mean that the BN is part of a specific family of graphical models known as directed acyclic graphs (DAGs). The majority of BNs are not time dependent, i.e., are 'static' in time, but nodes representing past events can be included and the requirement to incorporate time as a variable can be met with object-oriented and dynamic BNs (Johnson et al., 2010, Johnson and Mengersen, 2011). An example of a simple BN indicating factors influencing microbial growth is shown in Figure 3.2.

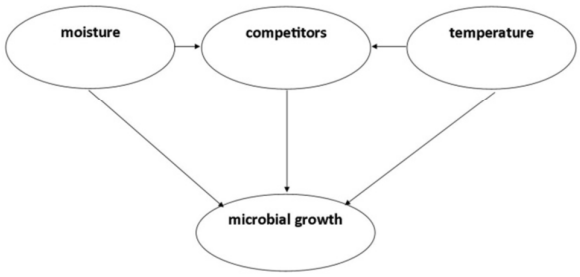

Figure 3.2. Example of simple Bayesian network indicating causal factors for microbial growth.

By their construction, BNs are able to characterise and quantify a complex outcome, as well as well as describe the many possible interactions between variables associated with the outcome (Donald et al., 2009). BNs can be used for 'forward inference', by which the inputs are specified and the impact on the outcome is observed and 'backwards reasoning', by which the outcome is specified and the states of the system's variables required to obtain that outcome are calculated. This

ability to undertake 'backwards reasoning' is due to Bayes' theorem (Jensen and Nielsen, 2007, Pearl, 2000). BNs are thus able to reveal variables that are major drivers for an outcome, or conversely, the sensitivity of the outcome to variables in the network (Ben-Gal, 2007, Coupé et al., 2000, Pollino and Hart, 2005). Uncertainty is explicitly represented in a BN, as each node or variable is represented as a probability distribution. This is particularly important in environmental systems, where uncertainty can be widespread (Aguilera et al., 2011).

Another useful feature of BNs is the faculty for 'structure learning', or the automatic derivation of the graph structure, either whole or in part, directly from a data set. This ability for BN structure to be directly induced from data is well established and has considerably increased the potential applications of BNs (Aguilera et al., 2011). Uusitalo (2007) argues however, that BN structures cannot be reliably estimated based on data as environmental systems include significant uncertainty and variability and further, that postulated beliefs about causal connections generally produce better models.

BNs have been used in environmental modelling for some years (Varis and Kuikka, 1999), although the full potential of these models in this field is thought to be largely untapped (Aguilera et al., 2011). In particular, they are only just emerging as popular models for describing the complex relationships between environmental exposures and health. Düspohl et al. (2012) maintain 'BNs have the potential to become a core method of transdisciplinary research and knowledge integration in environmental management'.

3.2.3 Bayesian networks in QMRA

BNs were proposed for QMRA over a decade ago (Barker, 2004). Their appearance in the field began and has predominated in food safety risk assessment, where they have been used to model one component of a QMRA (Gronewold et al., 2008), or a whole food production chain (Greiner et al., 2013). A stochastic QMRA, consisting of a set of biological and/or process-related variables and the mathematical equations defining their dependencies may be considered as a BN (Rigaux et al., 2012a, Smid et al., 2010). In other applications of a BN, a QMRA may provide inputs into a BN, or a BN may be used to augment a QMRA (Donald et al., 2009). Greiner et al. (2013) affirm that an entire QMRA model can be formulated as a BN using the same mathematical equations as an MC model but implemented in

a network which includes the joint distribution of all variables in the model. In general, Bayesian methods provide a flexible and powerful approach to QMRA and risk assessment modelling, with the caveat that implementation may be more challenging in practice than MC modelling (Greiner et al., 2013).

Applications of BNs in a QMRA have been described by previous authors in the food safety domain (Greiner et al., 2013, Parsons et al., 2005, Smid et al., 2010), but to the best of our knowledge, this is the first review of their use in the QMRA context across domains. Further elaboration of the advantages and drawbacks of BNs in QMRA occurs in the Discussion.

3.3 METHOD

Published studies were identified in a search of the Web of Science, Scopus, PubMed, SpringerLink and Informit databases, focusing on the terms 'Bayesian network', 'Bayesian belief network', or 'Bayesian graphical model', used in conjunction with the terms 'QMRA', or 'microb*'. QMRAs for foodborne or waterborne pathogens where BNs were used, with or without the use of Bayesian statistical methods, were included in this review. The literature describing applications of BNs in QMRA is not extensive, demonstrating the novelty of the approach in this particular domain. Fifteen papers were selected for inclusion based on their relevance to these search criteria. The studies were examined firstly to determine the study domain, aims and application of the method to QMRA. Closer scrutiny was undertaken to determine the knowledge source/s for the BN model structure, source of conditional probability table values, techniques used for validating the model and for belief updating (sometimes referred to as probabilistic inference). Finally, the identified functions of the BN (for example, prediction, separation of uncertainty and variability, scenario assessment and decision making) and other gains made in the course of the analysis e.g., software or new method development were ascertained. This information has been summarised in Table B1 in Appendix B.

3.4 RESULTS

Of the 15 peer-reviewed journal papers examined, 10 were published within the last 5 years. Eleven articles pertained to microbial risk assessment of foodstuffs and the remaining 4 related to waterborne microbial hazards. The number of nodes in

the BNs varied widely, from 6 to 63. In the following synopsis, each of these articles is now introduced and briefly discussed. The issues that they raise, in particular the strengths and drawbacks of the BN approach for QMRA, are then drawn together in a summary discussion.

3.4.1 Foodborne microbial risk assessments

The primary purpose of a risk assessment by G. C. Barker et al. (2002) was the translation of the QMRA to risk management decisions. A model of a nonspecific food manufacturing process was developed to represent two components of foodborne botulism (spore concentration and bacterial growth), in terms of contamination processes, spore thermal death kinetics, germination and growth of cells, toxin production and patterns of consumer behaviour. A BN was used to include such diverse information sources as operating experience, with low quality experimental data such as zero spore counts.

Microbial growth variability and uncertainty is a key source of microbial risk variability and uncertainty, which led Pouillot et al. (2003) to propose a method to estimate growth curve parameters of *Listeria monocytogenes* in milk, using published data. The primary aim of this BN application was to model separately and evaluate uncertainty and variability by means of hyperparameters, improving the growth model parameter estimation for risk assessment purposes.

In a QMRA quantified by a review of the scientific literature and industry practices, Parsons et al. (2005) used a BN to make estimates of the prevalence of *Salmonella*-positive birds in a flock and inferences about system variables, in order to reduce the *Salmonella* contamination rate in the final product of a poultry production chain. The QMRA was subsequently used as a basis on which to compare three modelling approaches for quantitative risk assessment.

Delignette-Muller et al. (2006) modelled the effects of time and temperature on competing growth rates of *Listeria monocytogenes* and food flora, as part of a larger collaborative project assessing exposure to the pathogen in cold-smoked salmon. The BN in this case accounted for the main sources of variability and uncertainty in these predictive microbiology models, thereby increasing the accuracy and validity of the models for the QMRA.

Albert et al. (2008) estimated the probability of contracting campylobacteriosis as a result of broiler contamination in a food production chain using a BN. In this instance a core stochastic model based on current or prior knowledge was built using only expert opinions and scientific literature. After an initial validation, the model was augmented with relevant data where it was available. The model illustrated the power of the Bayesian approach, particularly against a background of scarce data, as it enabled the combination of data with other disparate sources of information.

Articles by J. H. Smid et al. (2011) and J. H. Smid et al. (2012) describe the development and use of a BN to trace sources of contamination (biotracing) for individual *Salmonella*-positive pig carcasses in a slaughterhouse. The purpose of the model was to allow plant operators to prioritise decontamination measures. To achieve biotracing, a model must be able to answer questions in the reverse direction of the chain processing order, which requires the incorporation of multiple pieces of evidence to update the statistics of the model parameters. BNs allow for such inferential queries and are therefore an appropriate choice of model for biotracing, due to their ability to use downstream information to point to materials, processes, or actions within a particular food chain that can be identified as the contamination source. This model demonstrates the concept of biotracing, gives insight into the dynamics of *Salmonella* in the slaughter line and indicates where in the line data collection is most effective for biotracing.

In a QMRA undertaken by Rigaux, Ancelet, et al. (2012), genetic diversity and the variation in concentrations of *Bacillus cereus* with time and temperature in a processing chain for courgette puree were studied. The BN modelled batch-specific variability separately from uncertainty and enabled backward calculation to update the experts' knowledge about the microbial dynamics of the pathogen using experimental data. The results included improvement of prior beliefs about the dynamics of the foodborne pathogen and reduction in uncertainty.

Meta-analyses are increasingly being performed in QMRAs for food safety and quality, to estimate the inactivation or growth parameters of micro-organisms of concern, in order to generate sufficiently generic parameters, with their variability, which can be used in further quantitative risk assessments. Rigaux, Denis, et al. (2012) employed a BN to address a persistent problem in canned food processing, microbial spoilage by *Geobacillus stearothermophilus* (Rigaux et al., 2012b). The

BN was used to estimate the thermal inactivation parameters of the pathogen, using a meta-analysis of reference inactivation parameters for the organism, to take advantage of the large quantity of data in the scientific and grey literature.

J. Smid et al. (2013) used a BN to obtain an accurate estimate for the transfer ratio of bacteria from one surface to another during pork cutting in a processing chain, by incorporating uncertainty from one experiment and variability from multiple experiments into one model (Smid et al., 2013). Benefits included improved insight into biological parameters such as recovery ratios, pathogen count data and transfer ratios and a correct representation of their uncertainty, producing better QMRA models. The researchers attested that current approaches, in which uncertainty originating from limited count data is often neglected, lead to inconsistencies and an underestimation of the total uncertainty in a model.

A key driver for innovation in the UK dairy sector is the ability to deal rapidly with zoonotic hazards due to negative publicity. G. Barker and Gomez-Tome (2013) modelled *Staphylococcus aureus* in milk in terms of pathogen concentrations, population growth and enterotoxin production, as well as effects of cooling and storage on growth and alkaline phosphatase as an indicator of potential hazards. This BN also enabled food chain biotracing, indicating three potential causes of *S. aureus* contamination by propagating effects of particular end point observations to express posterior beliefs about possible causes.

3.4.2 Waterborne microbial risk assessments

Donald et al. (2009) developed a BN as a supplementary analysis to a QMRA, which described a conceptual model for health risks associated with recycled water, with a chosen health endpoint of gastroenteritis. The BN was useful in identifying the nodes with the most influence on the incidence of gastroenteritis. By calculating credible intervals the authors contributed a method for quantifying the uncertainty of point estimates arising from the BN, adding to the available tools for assessing microbial health risk as a result of environmental exposures.

Goulding et al. (2012) used a BN to increase understanding of the public health impacts of sewer overflows in wet weather, in order to prioritise management options. QMRA was used to identify the threats to the waterway values and the relationships between the variables for inclusion in the BN. The network model

enabled the effectiveness of various sewer overflow management options in reducing the public health risk to be determined through the application of probabilistic inference and the model was also able to account for the uncertainty inherent in such events and their subsequent impacts.

In a QMRA for waterborne pathogens in a freshwater lake, a Bayesian model was developed using concentrations of faecal indicator bacteria (FIB), frequency of pathogen detection and physicochemical parameters such as temperature and salinity to determine factors predictive of human health risk (Staley et al., 2012). The authors concluded that BN modelling of physical and bacterial parameters can be useful in predicting conditions under which low or high risk of pathogen presence exists, making the tool valuable in applications such as water quality monitoring at beach and shell fishing areas.

In a similar recreational water setting, the potential threat of faecal contamination was assessed using a BN to explore differences between analytical methods (most probable number (MPN) and colony-forming units (CFU) for quantifying FIB concentrations and between different sampling locations and times (Gronewold et al., 2011). The aim was to reduce uncertainty in water resource management decisions by fully understanding and accounting for methodological variability associated with FIB quantification methods and to improve the estimation and representation of FIB inactivation rates. Comparison of a conventional model of bacterial inactivation rates with a novel Bayesian model revealed that the latter provided a more robust approach to quantifying uncertainty in microbiological assessments of water quality than the conventional MPN-based model and therefore reduced uncertainty in water resource management decisions.

3.5 DISCUSSION

The synopsis above, along with the summary presented in Table B1 in Appendix B, clearly demonstrates the diversity and utility of BNs in the exploration of microbial risk. In general, the research on foodborne pathogens aimed to solve highly specific tactical or operational types of decision problems over short term time scales (Sutherland, 1983). In contrast, the research environments for waterborne pathogens were spatially larger, used aggregated state indicators and aimed to solve directive or strategic types of decision problems over longer time scales.

3.5.1 QMRA focus

In all of the studies, BNs were used in the context of QMRA to achieve the aim of quantifying an aspect of the microbial hazard and making predictions, although minimising risk through scenario assessment and informing management options was described in only 10 articles. Seven of the 15 studies reported using the method for the separation of uncertainty and variability and 8 mentioned reduction of uncertainty as a benefit of using a BN. Six of the 15 articles reported developing a new method or new software during the course of the research.

In the majority (11) of the 15 studies, the BN was used to investigate a fragment of the exposure pathway, such as the environmental influences on pathogen concentration, parameters of bacterial growth models, or pathogen enumeration issues such as recovery efficiencies. In the publications which did not describe a complete QMRA, it was often not clear whether the published material describing the BN application comprised the entire risk assessment. Studies by Donald et al. (2009) and Albert et al. (2008) incorporated the hazard identification, dose-response, exposure and risk characterisation steps of a QMRA. In one article (Goulding et al. (2012), a QMRA was used to provide inputs for their developed BN.

3.5.2 Foodborne risk assessments

Although most of the QMRA studies modelled specific modules of the food chain and thereby a component of the QMRA process in detail, Albert et al. (2008) began with a simplistic model of the entire food chain including consumption, improving certain points gradually with new evidence, which was subsequently propagated throughout the BN to maintain the overall veracity of the model.

The applications describing a fragment of the foodborne pathogen exposure pathway were very detailed, comprising extensive analyses of high resolution data. For instance, 8 of the 11 risk assessments of foodborne pathogens focused on the dynamics of microbial populations affecting an endpoint of pathogen or spore concentration, using a BN to improve estimates of variables such as growth and resistance parameters, or to examine the variation of populations with time and temperature. Another innovative purpose identified for BNs to augment QMRA is biotracing, the identification of sources of bacterial contamination in a chain of

events such as a food production line (Smid et al., 2011). Five of the 11 foodborne risk studies used a BN to achieve source-level inference, or biotracing.

3.5.3 Waterborne risk assessments

Of the four BNs used to assist in the description of waterborne pathogen risk, two chose pathogen presence/absence (Staley et al., 2012) or unobserved FIB concentration (Gronewold et al., 2013) as an endpoint. In a similar manner to the foodborne risk assessments with pathogen concentration endpoints, Staley et al. (2012) used a BN to undertake source tracking in order to identify faecal contamination sources in a freshwater lake and Gronewold et al. (2012) used a BN to explore the effect of the significance of sampling location and time on *Enterococcus* concentrations. The endpoint chosen by Donald et al. (2009) was gastroenteritis, whereas Goulding et al. (2012) expressed the risk to human health in terms of the threat to five waterway uses.

3.5.4 BN procedures

Of the 11 studies which stated their sources for the structure of the BN, 7 used a combination of sources and 4 used a single source to inform the model structure. There was a notable absence of accounts of structure learning from data, apart from that reported by Staley et al (2012). Equal numbers of studies (6) used empirical data or expert opinion to quantify conditional probability tables. Model validation was carried out principally using data or by sensitivity analysis, with 2 studies using existing models to validate their BN and 2 using expert evaluation. Validation procedures reported included conventional regression analysis, 10-fold cross validation and a ‘leave one out’ cross-confirmation procedure. In almost all cases, belief updating was achieved through the data, with 1 study also using expert opinion.

On the whole, explicit discussion of model validation, discretisation, belief updating and knowledge sources for conditional probability tables was uncommon and obscured by a lack of uniformity in terminology and structured detail. There was a wide variation in implied meanings of commonly used terms such as ‘Bayesian approach’, ‘Bayesian methodology’, ‘data’, ‘parameters’ and ‘variables’ and such terms were rarely defined by authors. Furthermore, while the literature was clear on modelling and statistical approaches, it was less clear on the mechanics of empirical

data being incorporated into conditional probability tables and on specific aspects of the application of the method in QMRA. The adoption of a standardised approach to the reporting of studies using BNs, as well as agreement on and use of a universal terminology would improve accessibility of this technique to multidisciplinary teams.

Based on the reviewed studies, primary advantages and drawbacks of the use of BNs for QMRA are now discussed.

3.5.5 Advantages of using Bayesian networks in quantitative microbial risk assessment

A BN is a natural framework for combining results of a QMRA with results from epidemiological studies. BNs facilitate the impartial, systematic combination of disparate information sources (Albert et al., 2008, Barker et al., 2002). Data can comprise point estimates, probability distributions, field observations, published results or expert opinion. Due to their ability to incorporate diverse data types, a BN enables a complex, multivariate statistical problem (such as QMRA) to be efficiently addressed where classical statistical methods are often inept (Albert et al., 2008). A particularly important feature with respect to QMRA is that poor quality experimental data has little impact on a distribution function for pathogen concentration that is established from high quality prior information (Barker et al., 2002, Kuikka et al., 1999). Furthermore, accurate predictions can be made with incomplete data (Fenton and Neil, 2013), or quite small sample sizes (Kontkanen et al., 1997).

As discussed previously, BNs by their nature allow 'backwards reasoning'. This means that when given evidence about an effect or outcome node, subsequent changes in the causal nodes can be observed. Standard MC approaches cannot use data sets downstream of other data sets (Albert et al., 2008). This feature enables determination of a diagnostic probability as opposed to a causal probability (Barker et al., 2009, Greiner et al., 2013). Moreover, new evidence can serve to update prior distributions, positively or negatively adjusting the initial beliefs of the expert (Greiner et al., 2013).

A BN responds immediately to changes in the network such as entering new evidence, because it does not use simulation. Information may be propagated from any point in the network to all others by Bayesian inference (Parsons et al., 2005).

The efficiency of a BN framework is also evident in the systematic representation of the joint probability in such a system, which significantly reduces complexity (Smid et al., 2010). Moreover, new evidence from multiple variables in the network can be used to update the estimates of the unobserved parameters in the model (Smid et al., 2010).

Interventions to reduce risk can be simulated in the network by changing parameter values, enabling calculation of risk reduction. An adequately validated BN provides a clear understanding of the effect of interventions on the outcomes of interest by altering the prior distributions of variables to simulate a new risk-reduction strategy (Albert et al., 2008). Verifying the effect of control measures by simulation greatly improves the visibility and efficiency of decision making, with the potential for reducing costs (Liu et al., 2013).

The visual representation in a BN of large quantities of complex information provides an informative platform for improved communications across disciplines between mathematical modellers, domain experts and stakeholders (Aguilera et al., 2011, Barker et al., 2002). Although their graphical structure is not a unique feature, environmental and biotic causal influences on the outcome of interest can be clearly represented and easily visualized (Smid et al., 2010). This feature is particularly relevant in QMRA methodology, where representation of a complex system is required in conjunction with transparent modelling of such process influences, as many of the leading-edge numerical tools employed in QMRA may not be accessible or transparent to non-technical team members (Smid et al., 2011).

3.5.6 Challenges of using Bayesian networks

Quantification of the conditional probability tables underlying each node in a BN can be challenging. Empirical data, from field or laboratory observations or from the relevant literature, require manipulation to determine conditional probabilities (Fenton and Neil, 2013). The alternative use of information elicited from experts to determine these conditional probabilities may also present a significant challenge (Düspohl et al., 2012, Newton, 2009), including representing these initial beliefs as a joint prior distribution over all the nodes in the BN, incorporating sufficient transparency and rigour in the elicitation process (Pollino and Hart, 2005) and overcoming the convictions of scientists to work with beliefs and probabilities when

they are familiar with observational data and classic statistical methods (Düspohl et al., 2012, Uusitalo, 2007).

A second challenge is the representation of probability distributions, either as continuous or discretised distributions. The discretisation of continuous variables introduces errors in the marginal distributions which can accumulate, producing an inaccurate representation of the model distributions (Parsons et al., 2005, Smid et al., 2010). Smid et al. (2010) assert that the discretisation of continuous variables leads to less accurate models, particularly when the distribution tails are critical (as in the case of low pathogen concentrations), as these are not explored in detail. However, an important development in the advancement of BNs has been the introduction of hybrid models, in which continuous and discrete variables can coexist (Aguilera et al., 2011). For continuous distributions, another limitation in the Bayesian approach is that the assumed distributional form of the priors may not be appropriate (Mitchell-Blackwood et al., 2012). The problem of subjectivity in model and parameter selection has also been identified as a drawback by Soller (2008).

The inability of BNs to support feedback loops due to their acyclic nature is a concern mentioned frequently in the literature (Jensen, 2001, McCann et al., 2006, Nyberg et al., 2006). This issue can be surmounted by the implementation of a dynamic BN (Johnson et al., 2010, Johnson and Mengersen, 2011, Smid et al., 2010), although Smid et al. (2011) caution that these models can become so large as to be computationally infeasible.

The choice of software used for the BN will determine whether certain procedures can be performed or not; for example, some software packages allow hybrid models, comprising continuous and discrete variables and some do not. Likewise, certain types of model validation, e.g., cross validation, can only be performed if the software accommodates it (Aguilera et al., 2011). Aguilera et al. (2011) also note that if the research is interdisciplinary, with both environmental and computer/mathematical roles, these software limitations may not pose a problem.

3.6 CONCLUSIONS

This appraisal of the use of BNs in assessing risk from microbial exposure provides a contextual background for the consideration of future frameworks and methods in this area. BNs provide a suite of attributes, including flexibility, the

modular representation of a complex multivariate problem, the ability to integrate different forms of information and to account for uncertainty and variability and the ability to make inferences using 'downstream' data. These features simplify scenario analysis during risk assessment and enable adaptive management, while the convenient graphical interface promotes ownership and communication among stakeholders and multidisciplinary research teams.

The drawbacks of BNs, including challenges with eliciting conditional probabilities and representation of spatial and temporal variability may depend on the complexity and scope of the risk question and can be overcome at least in part, by using them as an adjunctive modelling tool. The published research on microbial risk assessments with BN applications would benefit from increased emphasis on procedural transparency, organisational rigour and the inclusion of supplementary material with the primary article. We consider BNs have great potential for wider use in QMRA, for the protection of human health.

Chapter 4: Modelling microbial health risk of wastewater reuse: A systems perspective

Preamble

This chapter has been written as a journal article to meet Objective 2 of the research, as stated in the Introduction:

Objective 2 - To develop a conceptual model of influences on microbial health risk in a wastewater reuse context.

An unparameterised causal network or conceptual model of the system under study was considered an important precursor to the development of the BN models for characterisation and quantifying health risk. This phase of model development was essentially unconstrained, portraying significant causal influences on health risk.

This chapter is primarily my own work and the figures and tables were created by me. The article was published by Environmental International in November, 2015 and is reproduced here in its entirety. The reference for the publication associated with this chapter is:

Beaudequin, D., Harden, F., Roiko, A., Stratton, H., Lemckert, C., & Mengersen, K. (2015). Modelling microbial health risk of wastewater reuse: A systems perspective. Environment International, 84, 131-141.

Contributor	Statement of contribution
D. Beaudequin	Conception and conduct of research, model development, writing of manuscript, modifications to manuscript as suggested by co-authors and reviewers
F. Harden	Comments on manuscript, editing
A. Roiko	Comments on manuscript, editing
H. Stratton	Comments on manuscript
C. Lemckert	Comments on manuscript
K. Mengersen	Comments on manuscript, editing

Principal Supervisor Confirmation: I have sighted email or other correspondence for all co-authors confirming their authorship.

Name: Kerrie Mengersen Signature:  Date: 29/04/2016

Abstract

There is a widespread need for the use of quantitative microbial risk assessment (QMRA) to determine reclaimed water quality for specific uses, however neither faecal indicator levels nor pathogen concentrations alone are adequate for assessing exposure health risk. The aim of this study was to build a conceptual model representing factors contributing to the microbiological health risks of reusing water treated in maturation ponds. This paper describes the development of an unparameterised model that provides a visual representation of theoretical constructs and variables of interest. Information was collected from the peer-reviewed literature and through consultation with experts from regulatory authorities and academic disciplines. In this paper we explore how, considering microbial risk as a modular system, following the QMRA framework enables incorporation of the many factors influencing human exposure and dose response, to better characterise likely human health impacts. By using and expanding upon the QMRA framework we deliver new insights into this important field of environmental exposures. We present a conceptual model of health risk of microbial exposure which can be used for maturation ponds and, more importantly, as a generic tool to assess health risk in diverse wastewater reuse scenarios.

4.1 INTRODUCTION

Current predictions of the effects of climate change include worldwide water shortages (IPCC, 2014). As an estimated 80% of the world's wastewater is not collected or treated (UN-Water, 2014), recycling water that has previously been regarded as unusable serves a dual purpose, providing water for uses that are supplied unnecessarily by potable water and reducing the impacts of wastewater discharged into pristine or sensitive receiving environments (NRMMC-EPHC-AHMC, 2006). However concerns about health risk due to difficulties with detection and identification of pathogens, contribute to underutilisation of this resource. The systematic, comprehensive and transparent determination of the microbial safety of reclaimed water for specific purposes such as irrigation is a public health imperative when considering, for example, the estimated 3.5-4 million hectares in 50 countries that are irrigated with wastewater from varying sources (Haas et al., 2014).

Maturation ponds, a subgroup of waste stabilisation ponds, are used worldwide as secondary or tertiary wastewater treatment systems with the primary purpose of reducing the number of disease-causing microorganisms (Gloyna, 1971, Von Sperling, 2007). Maturation pond technology uses environmental influences such as sunlight (Maïga et al., 2009a) and pH (Curtis et al., 1992b) to inactivate pathogens and is used as a disinfection process for wastewaters in developing countries and in rural and remote locations in the developed world (Mara, 2004, Shilton, 2005). The treated wastewater progresses through the ponds over a period of several days to weeks in order to achieve pathogen reduction (Shilton, 2005). The effluent from these ponds may then be used for non-potable recycling purposes.

The aim of this study was to create a conceptual model of factors impacting health risk due to waterborne microbial exposures, integrating elements of the quantitative microbial risk assessment (QMRA) framework with environmental effects on pathogen concentrations. The study focused on sewage maturation ponds as a case in point. The objectives of the research were to inventory known and unknown factors influencing health risk in this scenario and to make assumptions explicit, thereby forming the basis of a future predictive health risk model for reuse of treated wastewater. To achieve the aim, the significant factors influencing health risk of maturation pond water reuse and their interactions, were identified and mapped using a participatory process. The result, an unparameterised causal network,

describes the significant influences on the health risk of an individual in a single exposure event, arising from exposure to the treated effluent at the point of discharge from a sewage maturation pond. This research is a component of a project with the objective of characterising and validating sewage maturation ponds with respect to pathogen removal with a view to reusing the treated wastewater, under the broader aim of increasing regional and remote water security in Australia.

'Systems thinking' (Capra et al., 2014, Meadows, 2008), based on the premise that the components of a complex environmental system are best studied from the perspective of their relationships with each other and other systems, was an underlying tenet of the research. Thus, human health risk was considered as an endpoint of a system, rather than the result of single factors such as pathogen concentration or dose, viewed in isolation. While this paper focuses on wastewater treated in maturation ponds, the QMRA submodels are generic and can be extrapolated to any water use or reuse options.

4.2 BACKGROUND

Microbial risk assessment of wastewater treated by a technology that is influenced predominantly by environmental conditions is made difficult by two main issues: the plethora of ecological influences on pond microbial populations and the complexity of human exposure pathways for diverse wastewater reuse scenarios. The common thread in these two issues is the myriad of interacting factors to be considered simultaneously to fully understand the health effects on the exposed individual or population. Robust inference in the public health interest in this context is difficult, as researchers attempt to quantify relationships in a virtually unbounded set of possibly correlated ecological variables, often with limited field data (Marcot et al., 2006). Development of a conceptual model, in which an archetypal set of key variables and their interactions is assembled and documented, provides a basis for communicating assumptions and completing a quantitative risk assessment (Suter, 1999) and is an important precursor to a sound predictive model.

4.2.1 Microbial health risk assessment

Risk assessment is the quantitative or qualitative determination of adverse consequences resulting from exposure to a hazard (enHealth, 2012). Quantitative risk assessment entails consideration of the magnitude of the adverse outcome and the

probability of its occurrence. The science of quantitative risk assessment is becoming increasingly complex. Improved research methods generate a plethora of data, resulting in multifaceted assessment of multiple risks and risks in vulnerable subpopulations (NRC, 2009). Risk assessment in its simplest form consists of four fundamental steps, illustrated in Figure 4.1. These are 1) hazard identification, 2) exposure assessment, 3) dose-response assessment and 4) risk characterisation (NRC, 1983). The framework has been adopted by many agencies worldwide and has since been further expanded to include risk management and risk communication steps (NRC, 2009).

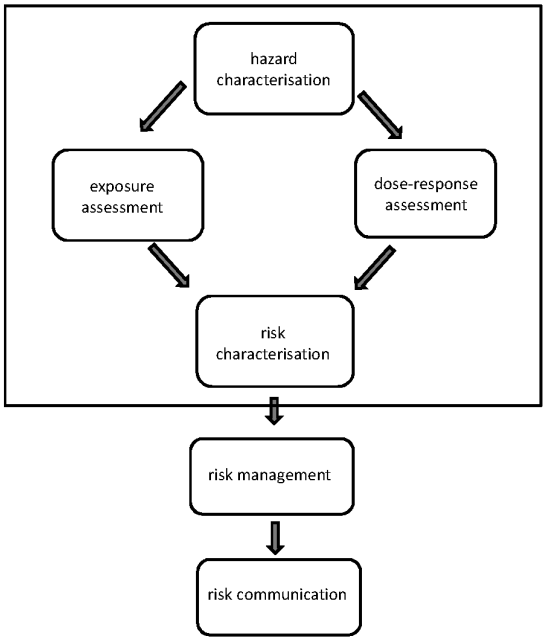

Figure 4.1. Generic risk assessment framework with original four steps (National Research Council, 1983) framed.

QMRA is a modelling process based on the National Research Council framework (1983, 2009) that is used for assessing public health risks that arise from exposure to pathogenic microbes, typically via food and/or water. This structured approach integrates information and data with mathematical models to examine the exposure and spread of microbial agents and thereby characterise the nature of adverse outcomes (CAMRA, 2013b, Havelaar, 2012, USEPA-USDA/FSIS, 2012). QMRA models can be used to generate knowledge about propagation of microbiological hazards along the risk pathway from source to exposure and effects

in complex real-world scenarios. This allows insight into the dependent relationships between input and output variables and the ability to quantify the effect of mitigation alternatives (Greiner et al., 2013). Haas (2002) envisaged that QMRA would eventually replace the use of indicator-based approaches to regulation of water quality (Haas, 2002). This has occurred for drinking water regulation in countries such as the Netherlands and Canada (Smeets, 2013) and for recycled water guidelines in Australia (Bichai and Smeets, 2013).

Current approaches to health risk analysis of water quality are constrained by a lack of empirical data and are challenging due to the high level of complexity inherent in natural systems. The process of empirical evaluation of health risk from exposure to pathogenic microorganisms, based upon the original four step risk assessment framework (NRC, 1983), has been broadened to account for the dynamic and epidemiologic features of diseases resulting from microbial infection (Fewtrell et al., 2001). The unique features of a dynamic infectious disease process not accounted for in the original risk framework, include: microbial growth and death; host immunity and susceptibility; the potential for secondary transmission; a range of possible health endpoints including delayed and/or chronic health effects; genetic diversity of microbial strains and their responses to interventions; detection method sensitivity; and multiple and sequential routes of exposure. Population, community and ecosystem dynamics, as well as heterogeneous spatial and temporal distribution in the environment, are other features of risk analysis involving microbial processes and systems requiring consideration in an idealistic conceptualisation of the pathogen-host-environment interaction (USEPA-USDA/FSIS, 2012)

4.2.2 Conceptual models

Conceptualisation of the risk pathway, beginning at the source of the hazard and ending at the significant undesirable consequences, has been described as the ‘backbone’ of every microbial risk model (Smid et al., 2010). Natural resource management often requires the representation of complex combinations of environmental, social and/or economic issues with uncertain outcomes, characterised by interactions across spatial and temporal scales, often in the absence of high quality observed data (Jakeman et al., 2006). Conceptual models, also described as unparameterised causal networks (Pollino et al., 2007), or ecological causal webs (Marcot et al., 2006), wherein variables and their relationships are advanced by

experts but not validated with data, are therefore invaluable tools commonly used in environmental domains (Low Choy et al., 2009). The development of the conceptual model can be regarded as a qualitative analysis of the system or problem, wherein specific expertise is sought from experts, partners and consultants regarding steps, processes and variables of influence in the various domains in the model.

After definition of the problem, the initial modelling phase may be achieved by a preliminary review of the literature and consultation with domain experts. The key variables in the system and influential parameters are mapped in a causal network (Marcot et al., 2006), that can be iteratively updated via a participatory learning process involving the modeler, multidisciplinary stakeholders and domain experts (Barton et al., 2012, Jakeman et al., 2006). The modelling process can define the scope of the research, make assumptions explicit and reveal their implications, inventory what is known and what is not, explore possible obscure outcomes and appraise the impact of changes and interactions on outcomes. An appreciable benefit of the process is the enhancement of communication between researchers from different backgrounds and between researchers and the broader community (Jakeman et al., 2006). As the conceptual model becomes more sophisticated there is a reduction of uncertainty (Thoeye et al., 2003), however with enhancement of precision there is a concomitant requirement for more data and transparency of results may be lost (Zwietering, 2009).

4.3 METHOD

Identification of the principal influential factors and development of the model structure was achieved iteratively through a series of meetings with domain experts. In the first phase, a meeting of project partners and consultants to the project took place, comprising representatives from the water industry and state health authority and researchers from the disciplines of microbiology, ecotoxicology, hydrodynamics and health risk modelling. The purpose and boundaries of the model and scope of the health risk assessment were determined at this initial meeting. The first version of the model, based on the four steps of the risk assessment framework (NRC, 1983) was then constructed from the literature on microbiological risk assessment methods and sewage maturation pond operation and performance. Major variables of influence and their interactions were identified in peer-reviewed journal articles and seminal texts and entered as nodes in a directed acyclic graph (Korb and Nicholson, 2011),

with arrows between the nodes representing causal links. The scope of this phase of the modelling process was constrained only by the model boundaries established at the initial meeting. Both known and hypothetical factors of influence in the system and their interdependencies were included in this phase of model development.

In the second phase, the model was reviewed for errors and omissions by a subgroup of the project team, comprising the research scientists representing the disciplines enumerated previously. Participants were given hard copies of the model to consider and were asked for their feedback. As a result of this step, a small number of additional variables were proposed, after which it was agreed that all of the important variables in the system had been captured. The model was then presented to and critically evaluated by an academic audience and an independent microbial risk consultant, none of whom had been included in previous deliberations. The final version of the network was presented at a full meeting of the project team and the model was endorsed as an accurate representation of a generic maturation pond system and microbiological risk assessment process. The model is described in detail in section 4.4 Results.

4.4 RESULTS

The conceptual model comprises four submodels (Figure 4.2), each of which will be discussed in detail in subsequent paragraphs. The four submodels are: *Pond operation and performance* submodel, representing key influences on the concentration of pathogens in a maturation pond system; *Exposure* submodel, incorporating factors to be considered in the characterisation of exposure to pathogens; *Dose-response* submodel, incorporating factors to be considered in the characterisation of the dose-response relationship and *Risk characterisation* submodel, representing combination of the *Exposure* and the *Dose-response* models and considering the disease outcomes to be considered in estimation of health risk. The submodels are linked when outputs of one submodel become inputs to another submodel.

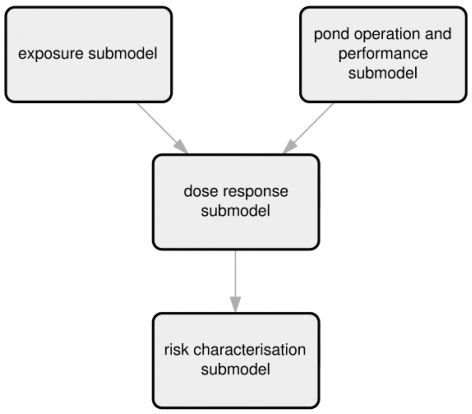

Figure 4.2. Conceptual model for assessment of microbial health risk associated with exposure to wastewater treated in a maturation pond.

4.4.1 *Pond operation and performance* submodel

The *Pond operation and performance* submodel (Figure 4.3) represents the factors influencing the pathogen concentration in the finished effluent at the end of the pond treatment process.

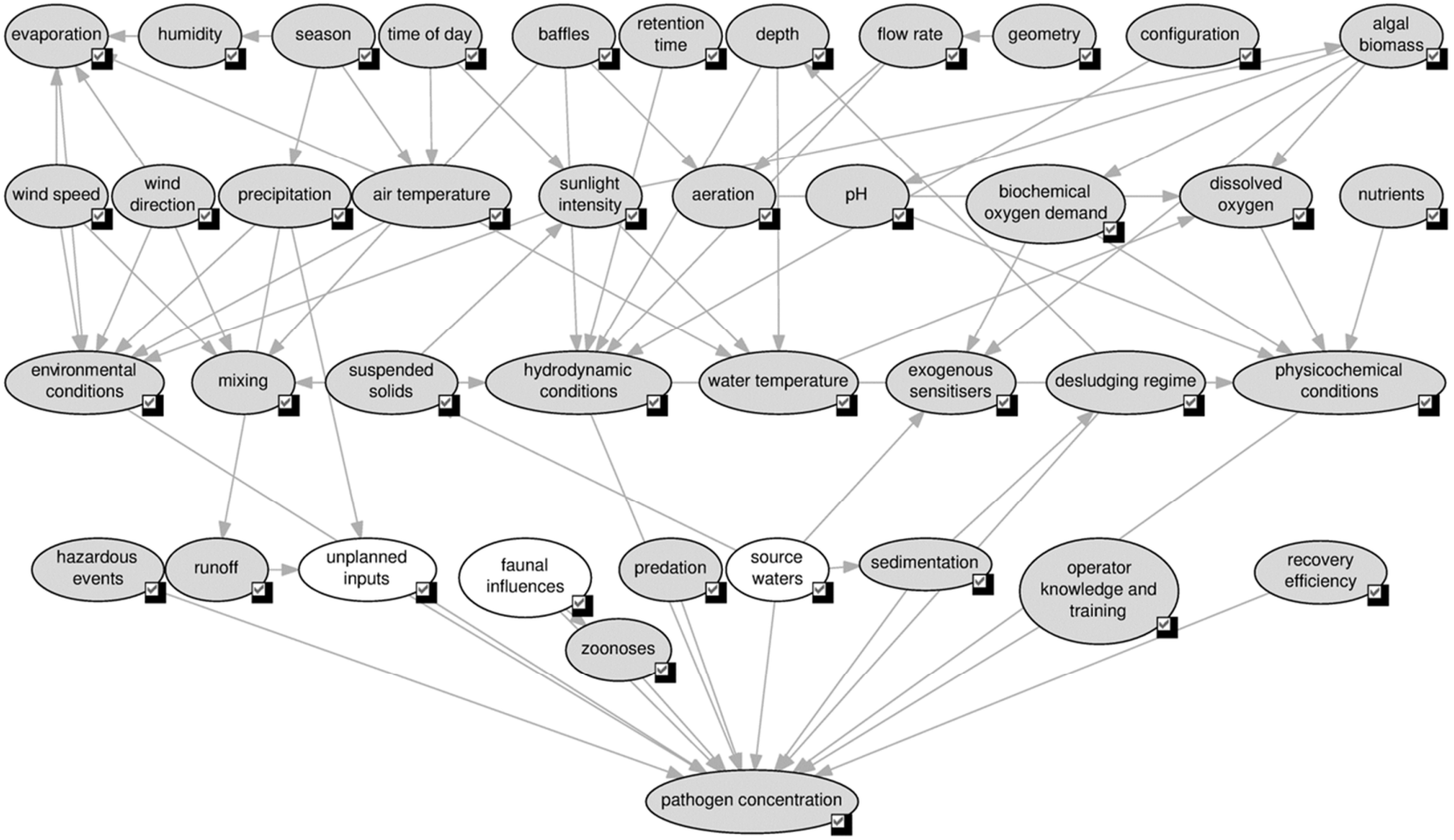

Figure 4.3. Conceptual model of factors influencing pathogen concentration in a sewage maturation pond with pathogen sources indicated as unshaded node

Pond performance

Maturation ponds are low cost wastewater treatment systems (usually 1-1.5 m in depth), that function most efficiently in warm climates (Mara, 2003). They are often found in series with anaerobic and facultative ponds. Their principal role is destruction of enteric pathogens (Mara, 2003) and when functioning optimally, they are capable of removing 90-99% of bacterial pathogens (Bitton, 2005, Von Sperling, 2007). In warm climates, with ambient temperatures exceeding 20°C, a waste stabilisation pond system with 4-5 ponds in series and a 20-30 day retention time is capable of reducing faecal coliforms by 4-6 log units. The same pond system can reduce enteric viruses by 2-4 log units, mostly remove helminth eggs and reduce biochemical oxygen demand by about 80% (Shuva and Fattal, 2003). The major physicochemical and environmental factors influencing the performance of a maturation pond are light intensity, pH, dissolved oxygen, wind and temperature (Sah et al., 2012).

Internal biochemical processes in ponds

Despite the apparent simplicity of the treatment pond concept, pond treatment processes are still not entirely understood, due to the large number of factors involved, their interplay and temporal and spatial variation (Sah et al., 2012). For example, sunlight has been shown to be a key factor in the inactivation of faecal indicators (Curtis et al., 1992a, Davies-Colley et al., 2000, Maïga et al., 2009a)and pathogenic bacteria (Boyle et al., 2008), but it can also be a temperature-dependent process (Maïga et al., 2009a), that is influenced by physicochemical factors such as dissolved oxygen (Jori and Brown, 2004), pH (Curtis et al., 1992b, Davies-Colley et al., 1999) and depth (Maïga et al., 2009b). As microbial inactivation continues in dark conditions (Craggs et al., 2004), mechanisms other than light are also thought to contribute to disinfection in ponds. Inactivation of microorganisms by sunlight can be further enhanced by exogenous photo-sensitising substances in pond water such as humic acids or algal compounds that promote light absorption and also by endogenous cellular photo-sensitisers (Curtis et al., 1992b). Furthermore, the mechanism of effect may differ between viruses, bacteria and parasitic pathogens (Sinton et al., 2002) and between species in the same pathogen class, as outlined below (Kadir and Nelson, 2014).

In maturation and other types of oxidation ponds, heterotrophic bacteria and algae exist in a symbiotic relationship, whereby the algae generate oxygen through photosynthesis and the bacteria use the oxygen to break down organic material, producing carbon dioxide that is fixed into carbohydrates by the algae (Mara, 2003). Algal photosynthesis increases pH, which is thought to contribute to pathogen destruction (Bolton et al., 2010, Curtis et al., 1992b) particularly at values over 9 (Pearson et al., 1987). Faecal bacterial removal rates are also proportional to temperature and retention time, but are inversely proportional to biochemical oxygen demand and pond depth (Saqqar and Pescod, 1992). Other factors influencing destruction of bacterial pathogens include predation by zooplankton (Bitton, 2005), aeration, nutrient depletion and sunlight intensity (Fernandez et al., 1992, Qin et al., 1991). Enteric viruses are also thought to be inactivated in maturation ponds by high temperatures, intense solar radiation and high pH (Bitton, 2005). Viruses can adsorb to settleable solids including algae and be removed from the water column by sedimentation (Mara, 2003), however they may survive for longer periods in the pond sediments than in the water column (Bitton, 2005). Efficacy of removal of helminth eggs and protozoan cysts is influenced by pond retention time, temperature, pH and solar radiation (Bitton, 2005). Sedimentation has been reported to be a significant factor in parasite removal (Mara, 2003, Von Sperling, 2007) but this has been debated by some authors (Reinoso et al., 2011). Contrary to previous studies, Reinoso et al. (2011) demonstrated that physicochemical factors (light, pH, dissolved oxygen, ammonia concentration) can be the primary cause of the removal of parasites from these systems and that sedimentation as a removal mechanism was less important than had previously been estimated.

Hydrodynamic considerations in ponds

In addition to internal biological and physicochemical processes, treatment efficacy is also strongly influenced by hydraulic conditions (Moreno, 1990). Retention time is a key factor, since the internal biochemical processes require time to achieve disinfection of the raw wastewaters (Lloyd et al., 2003, Vorkas, 1999). Mixing is another important aspect of pond dynamics and is influenced by variations in water temperature stratification and by wind conditions (Brissaud et al., 2003) and the presence of flow directing vanes or panels, termed baffles (Olukanni and Ducoste, 2011). Other hydrodynamic influences on pond performance include the

geometric shape, specifically length to width ratio (Abbas et al., 2006, Olukanni and Ducoste, 2011) and the configuration of the pond system. Optimal disinfection efficiency is achieved with using one of two pond configurations: either a single pond with baffles (Mara, 2009) or channels (Bracho et al., 2006), or three or four ponds in series (Von Sperling, 2007). Operational factors affecting the ability of a pond to destroy pathogens include the desludging regime and operator knowledge and training.

External environmental factors that influence pond performance include unplanned inputs such as torrential runoff from surrounding terrain, faunal influences such as birds and turtles inhabiting the pond and changes in the characteristics of source waters. There may also be seasonal variations in human infections with pathogens such as *Cryptosporidium* and these may affect influent pathogen load due to increased shedding (Cunliffe, 2006). In addition to regular variations in disinfection efficacy due to environmental influences, hazardous events such as equipment or power failure, or heavy rainfall can result in short periods of reduced efficacy, contributing to peaks in pathogen concentration, potentially increasing health risk. An assessment of the frequency, duration and magnitude of hazardous events is essential in QMRAs of water treatment processes. Alternatively, they must be modelled separately in a dedicated QMRA for hazardous events (Smeets et al., 2006). The *Recovery efficiency* node (Figure 4.3) provides an evaluation of the accuracy of the pathogen enumeration method in estimating the true pathogen concentration in field and laboratory observations. It is widely acknowledged that pathogen enumeration data are inherently variable due to random errors in sample collection, processing and counting (Petterson et al., 2007, Schmidt et al., 2010), thus reducing the accuracy of concentration estimates. It has been suggested that system-specific recovery data, or at least estimates of the recovery fraction, be incorporated into concentration estimates so as not to underestimate the risk, especially when concentration estimates are used to infer human health risks (Petterson et al., 2007).

The outcome node of the *Pond operation and performance* submodel, *Pathogen concentration*, becomes an input to the *Pathogen dose* node in the *Dose-response* submodel (Figure 4.4), thereby linking these two submodels.

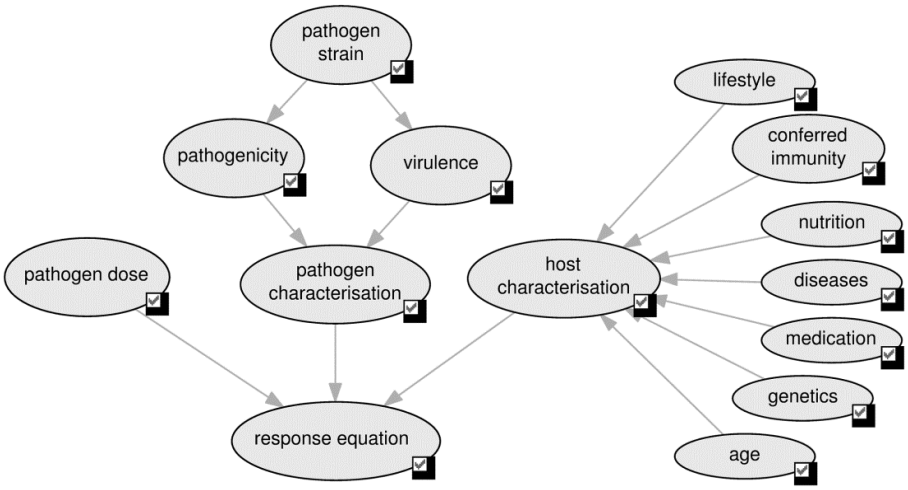

Figure 4.4. Dose-response sub-model – variables for consideration in modelling dose-response in a microbial risk assessment.

4.4.2 *Dose-response* submodel

The *Dose-response* submodel illustrated in Figure 4.4, describes the factors influencing individual response to pathogen dose. In the 'epidemiological triangle' illustrated in Figure 4.5, the dose-response step of a microbial risk assessment represents the interaction between the pathogen and the host. The outcome of the dose-response step is an estimate of the probability that an individual will exhibit a defined physiological response as a result of exposure to a stipulated dose of a specific pathogen. The three central nodes in the *Dose-response* submodel, *Pathogen characteristics, Pathogen dose and Host characteristics*, represent the myriad of factors influencing the inter- or intra-individual variability in the human response to a given dose of a pathogen.

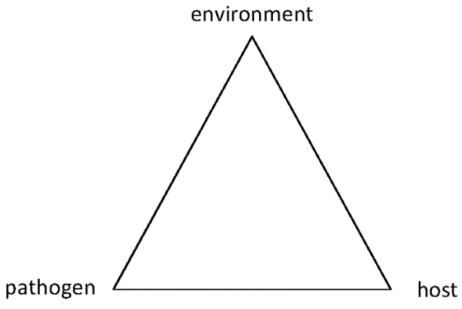

Figure 4.5. Epidemiological triangle.

Pathogen characterisation

The *Pathogen characterisation* node represents factors such as pathogen strain variability resulting from differences in genetic lineages and polymorphisms, resulting in differences in pathogenicity: the ability of the organism to cause disease and virulence: the degree of pathology the pathogen is capable of causing (Buchanan et al., 2000). The latter is usually correlated with the ability of the pathogen to multiply within the host (Buchanan et al., 2009).

Dose

The *Pathogen dose* node, representing the actual number of pathogens invading the host, is a key input variable for the R*esponse equation* node. Pathogen dose is calculated from the number of pathogens or infective particles in the medium and the volume of the medium implicated in the exposure. Experimental observations show that the probability of acute illness among infected subjects may increase with increased dose, but a decrease has also been demonstrated (Teunis et al., 1999).

Host characterisation

The *Host characterisation* node represents the factors that influence individual physiological response to pathogen dose. These include genetics (Buchanan et al., 2000, Zeise et al., 2013); age (Gerba et al., 1996, Nwachuku and Gerba, 2004, Teunis et al., 2002); pre-existing diseases that impair immunity such as HIV/AIDS, diabetes or cancer; nutritional status (Buchanan et al., 2000, Zeise et al., 2013); lifestyle factors such as cardiovascular fitness and substance use; previous exposure conferring immunity to the pathogen (Buchanan et al., 2000, Drechsel et al., 2010); and prescribed medicines (Buchanan et al., 2009, Juliens et al., 2009). Conferred immunity to the pathogen can fluctuate depending on the time since last exposure and the presence of concomitant infections (Buchanan et al., 2009, Juliens et al., 2009, USEPA-USDA/FSIS, 2012). Currently, microbial risk assessment models do not account for conferred immunity (Havelaar and Swart, 2014). However, in a recent campylobacteriosis case study, the standard approach to risk characterisation without accounting for conferred immunity and conditional probability of illness given infection, resulted in overestimation of incidence of disease by several orders of magnitude. An extension of current dose-response models to include these factors was proposed (Havelaar and Swart, 2014).

Other factors thought to influence the host response include medications affecting stomach pH (e.g., antacids, proton pump inhibitors) or those that alter gut transit time (e.g., opioids, laxatives), as these impair the first line of defence against an ingested pathogen. In addition, immuno-compromised status from autoimmune and immunodeficiency diseases, some cancers or pharmaceuticals such as nonsteroidal anti-inflammatory, cytotoxic or immunosuppressive agents, increase individual susceptibility to and severity of infection. It is clear that a great deal remains to be done in exploring and mapping the variability in human susceptibility to microbial infection. Refinement of predictive models requires increased understanding of the underlying biology, as well as further exploration and quantification of sources of variability in dose response (Buchanan et al., 2009).

Response equation

The *Response equation* node represents the pathogen-specific mathematical models that estimate the probability of a defined physiological response following exposure to a pathogen dose. A limited number of models have been developed and published, using data from outbreak or feeding studies on animals or human volunteers (CAMRA, 2013a). However uncertainty arises when generalising from experimental datasets on relatively homogenous, healthy test populations to realistically variable exposed populations (Buchanan et al., 2009). Challenges to dose-response modelling include: multiple dosing; interaction with *in vivo* processes; incorporation of host susceptibility factors; gauging the time to effect; variability among pathogen strains; route-to-route extrapolation; validation of animal models with outbreak data; and, of particular relevance to wastewater reuse, concomitant microbial and chemical exposures (Haas, 2011). The 'defined physiological response' to a pathogen is an important component of the dose response and varies widely among models. Examples of response definitions used by the Centre for Advancing Microbial Risk Assessment (2013a) include death, positive isolate in stools, corneal ulceration, shedding in faeces and stillbirths.

The outcome of the *Dose-response* submodel, *Probability of infection,* represents the probability of an individual becoming infected as a result of a single exposure event. It becomes an input node for the *Risk characterisation* submodel, linking the *Dose-response* and *Risk characterisation* submodels.

4.4.3 *Exposure* submodel

Figure 4.6 represents the *Exposure* submodel. The microbial exposure assessment or 'risk pathway' has been described as the foundation of every microbial risk model (Smid et al., 2010). It has the greatest degree of variability and uncertainty (Teunis et al., 1997) and is arguably the most difficult step in QMRA (Covello and Merkhofer, 1993). In the 'epidemiological triangle' (Figure 4.5), exposure can be considered as the interaction between host and environment and pathogen and environment. The nodes in the *Exposure* submodel describe factors in the context of a defined wastewater reuse that contribute to the primary exposure pathway of waterborne pathogens.

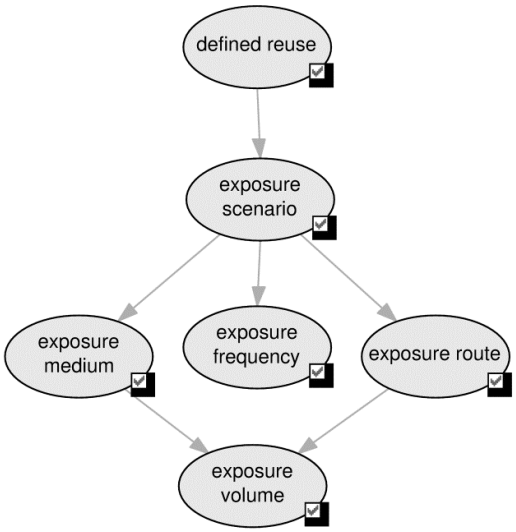

Figure 4.6. Exposure sub-model – variables for consideration in modelling exposure for a microbial risk assessment of wastewater reuse.

Designated wastewater reuse

The *Defined reuse* node designates the use of treated wastewater, such as irrigation of municipal areas, or irrigation of commercially grown food crops (NRMMC-EPHC-AHMC, 2006) and the *Exposure scenario* node describes the combination of circumstances in which contact occurs with treated wastewater. In these two reuse situations, there will be different pathways of exposure for a child playing in a public park in the former case, or a farm worker in the latter example. Consequently the reuse scenario under consideration will influence exposure medium and route, as well as frequency of exposure.

Exposure route

The *Exposure route* node represents the route of entry of the pathogen into the host. The way in which individuals are exposed to pathogens in treated wastewater will depend on how the water is used and the scenario under consideration. Some pathogens may have more than one route of transmission. For example, enteroviruses may be transmitted via ingestion or inhalation (Haas et al., 2014). The principal transmission pathways for pathogens in recycled water include: direct ingestion of contaminated water, droplets or airborne particles; direct ingestion of food that has been contaminated by pathogens from recycled water; indirect ingestion of pathogens via licking of fingers or objects that have touched contaminated surfaces; direct inhalation of contaminated water droplets and aerosols; and direct contact with skin, eyes or ears (EPA QLD, 2005).

Direct ingestion, such as through consumption of uncooked leafy vegetables irrigated with treated wastewater, is the most documented and studied route, potentially delivering the largest dose of pathogens and so having the greatest risk of causing disease (EPA QLD, 2005). Research on the effects of exposure through other routes is generally focused on specific pathogens, such as inhalation of *Legionella pneumophila* (Thoeye et al., 2003). Determination of the exposure route with the highest risk may vary depending on the reuse, as in inhalation of aerosolised wastewater in farm workers during spray irrigation (Thoeye et al., 2003). A QMRA can be based on the exposure route with the highest risk, or consider multiple exposure routes. The *Exposure route* under consideration is a determinant of *Exposure volume.*

Exposure medium

The *Exposure Medium* node represents the matrix, such as air, soil or food that conveys the pathogen to the host. Characteristics, including nature and consistency, are important components of an exposure assessment (O'Toole, 2011), as they influence infectivity, growth, decay and spatial distribution of pathogens. For example in a wastewater medium, microorganisms may embed themselves in, or clump around particulates such as algae or suspended solids. If pathogens are thus distributed heterogeneously in a delivery medium, the estimation of pathogen concentration and therefore dose from a given volume of the medium may vary widely between exposure events (USEPA-USDA/FSIS, 2012). If the concentration

of pathogens is very low, as in highly treated wastewater, exposures to the same volume of medium may result in a zero dose of pathogens, while other exposures deliver one or more pathogens. Characteristics such as pH and nutrient content may also influence microbial inactivation or growth (USEPA-USDA/FSIS, 2012). Therefore it is recommended that the delivery matrix being considered in the exposure scenario be comparable with that used in generating dose-response data (USEPA-USDA/FSIS, 2012), although to the author's knowledge this is currently not possible for wastewater exposures as no dose-response data on pathogens in wastewater exist.

Exposure frequency

Exposure frequency describes how often an exposure event takes place e.g., number of times per year (NRMMC-EPHC-AHMC, 2006) and contributes information for the production of a standardised risk estimate which can be compared with other such estimates. The exposure frequency for a defined reuse such as municipal irrigation may be influenced by factors such as whether the exposed individual at the facility is a municipal worker or member of the public and in the case of a commercially irrigated food crop, whether the exposed individual is a farm worker or a consumer.

Exposure volume

The *Exposure volume* represents the volumetric quantity of the exposure medium such as soil, air or water, during a single exposure event. In this model, exposure volume will be influenced by the designated reuse of the water, the reuse scenario under consideration and the exposure route. The *Exposure volume* node is a key variable, providing input for the *Dose* node in the *Dose-response* submodel, thereby linking the *Exposure* and *Dose-response* submodels.

4.4.4 *Risk characterisation* submodel

Dose-response models can be further considered as dose-infection and infection-illness models (FAO/WHO, 2003), although the preponderance of existing dose-response models reflect the dose-infection step. Infection-illness models, representing the proportion of infected individuals who develop symptoms of illness, are currently the exception rather than the rule and data available are limited (Havelaar and Swart, 2014). Other information currently lacking in illness models

includes incubation times, duration of illness and timing of immune response (FAO/WHO, 2003).

Manifestation of infection

The *Risk characterisation* submodel illustrated in Figure 4.7 represents the risk characterisation step in QMRA, where information from exposure and dose–response assessments is combined to portray outcomes of infection, indicating the frequency and severity of risk to health of populations. The input node to this submodel is the outcome node *Probability of infection* from the *Dose-response* submodel. Although a definition of infection has not been universally agreed upon (FAO/WHO, 2003), it is generally accepted that infection is defined as multiplication of organisms within the host, evidenced by measurable rises in serum antibodies and/or excretion of the organism, with or without clinical symptoms (Haas et al., 1999). Thus, since clinical infection may or may not result in the appearance of symptoms (Haas et al., 2014), it is further considered in this submodel as the nodes *Asymptomatic infection* and *Symptomatic infection*.

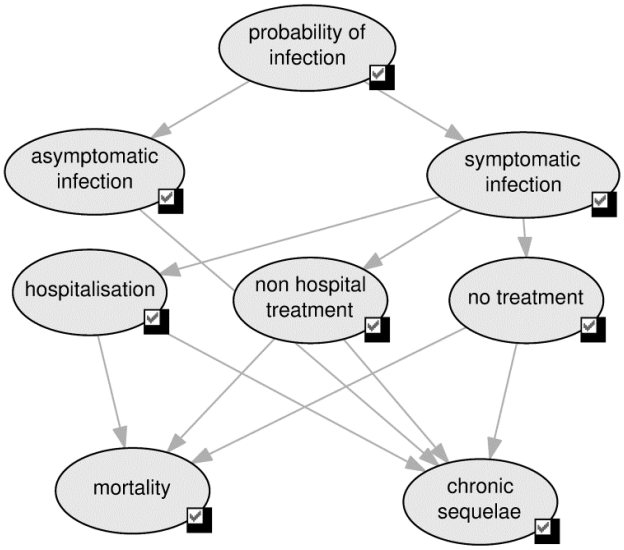

Figure 4.7. Risk characterisation sub-model, containing nodes which represent the outcomes of infection.

Level of treatment

The *No treatment*, *Non-hospital treatment* and *Hospital treatment* nodes delineate severity of illness outcomes. *No treatment* represents the proportion of infected individuals who self-treat or do nothing, *Non-hospital treatment* represents

those who access treatment at a general practice, pharmacy or other community-based practitioner and *Hospital treatment* represents the proportion of those who present at hospital emergency departments. The *Mortality* node represents the proportion of individuals who die as a result of acute infection.

Infection outcomes

The outcomes of pathogen infiltration typically modelled in QMRA are infection, illness and death (USEPA-USDA/FSIS, 2012). For some pathogens however, the cascade of consequences can be quite complex, involving multiple disease symptomatologies and endpoints. For example, symptoms of enterohaemorrhagic *Escherichia coli* range from relatively mild fever, vomiting and diarrhoea to haemolytic uraemic syndrome and potentially, stroke and renal failure (WHO, 2011). Portrayal of the true burden of disease is important in producing accurate cost estimates and informing decision making and policy development (Keithlin et al., 2014). Consequently the *Chronic sequelae* node represents delayed or secondary adverse health effects that occur as a result of a microbial infection with symptoms that differ from the initial acute reaction. Some secondary sequelae such as joint inflammation and reactive arthritis associated with infections from *Salmonellae*, *Shigella* spp. and *Campylobacter jejuni* become chronic (USEPA-USDA/FSIS). Both delayed and chronic sequelae may result from either asymptomatic or symptomatic infection.

The occurrence and frequency of severe disease outcomes is often better recognised and quantified than that of less severe outcomes (Haas et al., 2014). This is because asymptomatic infection or mild illness does not require medical care. In the case of mild symptoms, the causative agent may not be recognised (Haas et al., 1999). Furthermore, linking waterborne microbial infections with secondary complications is difficult as associative illnesses are not likely to be identified. Chronic effects may occur at a much later date and will also not be linked to the precipitating infection, as mild acute infections are rarely documented and acute illnesses are not typically followed over time to observe secondary or chronic effects (Lindsay, 1997).

The output of the *Probability of infection* node can be used, along with data on annual frequency of exposure and ratio of illness to infection, to calculate annual risks of infection and illness which can be included as nodes or submodels and

compared with established benchmarks. Similarly, since use of DALYs is the usual method for characterising and comparing health risks (WHO, 2008, Havelaar and Melse, 2003), a node representing DALYs can be added and/or the variables required to quantify same. The *Risk characterisation* submodel represents a parallel approach to annual risk estimates and DALYs for characterising health outcomes and employment of one approach does not preclude the other. Figure 4.8 shows the sewage pond operation and performance submodel, linked with the three submodels representing the QMRA process, to form the conceptual model describing factors influencing microbial health risk of treated wastewater.

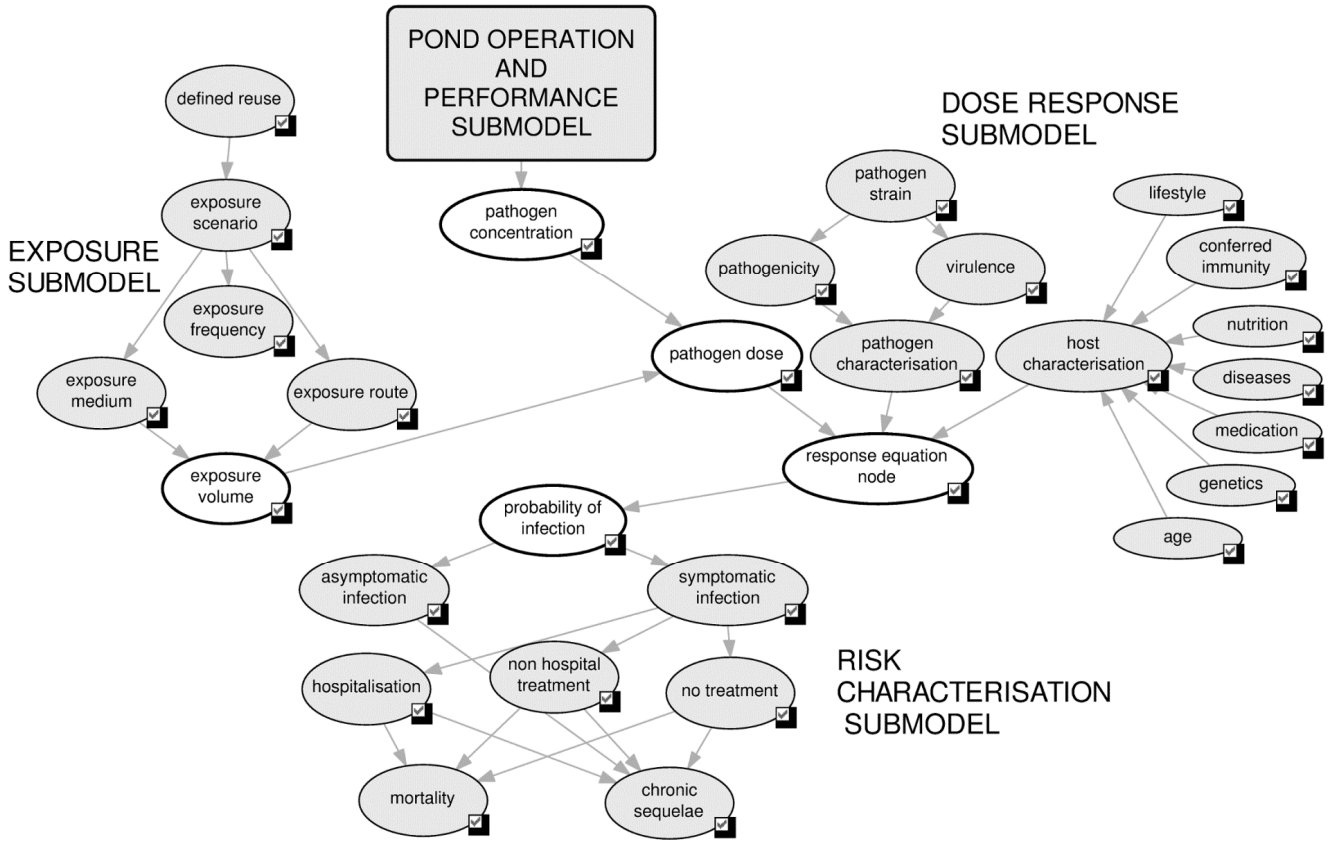

Figure 4.8. Conceptual model of wastewater reuse based on QMRA framework. Nodes linking submodels are unshaded.

4.5 DISCUSSION

The above overview demonstrates the complex interplay of factors involved in characterising health risk of exposures to water treated in maturation ponds. Use of a systems approach in modelling health risk is now illustrated and discussed.

4.5.1 QMRA purpose and the systems approach

The type of QMRA to be undertaken is determined by the scope of the problem, the goal or required outcome and available data. For example, a screening QMRA is a conservative estimate of possible risk based on available data, is usually simple and able to be achieved rapidly; a risk ranking QMRA may compare risk among several hazards, such as a single pathogen evaluated in multiple wastewater reuse scenarios, a single water source containing multiple pathogens, or multiple treatment types or reuse options; a product pathway assessment identifies the key factors affecting exposure including potential impact of mitigation strategies; a risk-risk analysis considers the trade-off of one risk for another and a geographic risk assessment examines the factors which either limit or enhance risk in a given region (USEPA-USDA/FSIS, 2012). Furthermore, researchers undertaking QMRA commonly focus on a limited number of components and highly specific questions regarding the infection pathway (Albert et al., 2008, Soller et al., 2004). A benefit of using a systems perspective in QMRA is that the model can be compartmentalised and generic, providing a common platform for describing a large range of problems of varying complexity (Beaudequin et al., 2015b). Scale and feasibility can be determined by availability of data, management or decision-making objectives and or biological relevance.

4.5.2 Static versus dynamic models

The model presented here is a static model, in that risk is characterised at an individual level (Eisenberg et al., 2002). Issues such as person-to-person transmission and immunity can be taken into account in a dynamic risk assessment, where the number of individuals assumed to be susceptible to infection varies with time and risk is assessed at the population level (USEPA-USDA/FSIS, 2012). Although dynamic models are capable of approximating biological ‘realism’, they may be analytically complex, resulting in increased computational demands and variability due to the uncertainties associated with model specification (Soller and

Eisenberg, 2008). Conversely, a simpler model involves implicit or explicit assumptions that may or may not be realistic or appropriate for a particular situation. Soller and Eisenberg (2008) argue in favour of parsimony, showing that a static model provides risk estimates of acceptable precision in low risk conditions (Soller and Eisenberg, 2008). Nevertheless, a temporal component can be added to the static systems model presented here, by incorporating a generic time slice that is then reused (Jensen and Nielsen, 2007, Johnson and Mengersen, 2011, Nicholson and Flores, 2011). Similarly, a spatial component can be added to produce a model that includes geographic information. Model nodes can be quantified using data from a geographic information system, or model output can be used to determine whether different processes or interactions apply for different locations (Barton et al., 2012, Johnson et al., 2012).

4.5.3 Sensitive populations

Sensitive populations with decreased immune capacity represent as much as 40% of the general population in the United States (Haas et al., 2014) and this proportion is expected to increase, due to increased longevity and the number of immunocompromised individuals. Sensitive populations include: the young and the elderly with immature and decreased immune responses, respectively; those taking drugs which suppress immune function or decrease body defences; pregnant women; and immunocompromised individuals with pre-existing autoimmune or immunodeficiency conditions. It is uncertain whether these groups are simply more susceptible to an initial infection and whether they are at increased risk from infection with particular pathogens, or from all pathogens. It is also unclear whether the infection rates for the immunocompromised and immunocompetent are similar, but the former group is more likely to become symptomatic. Whether these susceptible populations should be represented as a tail of the normal dose-response distribution or as a different population is also a matter for debate (Buchanan et al., 2000). A systems-based model such as that presented here, comprising common core nodes, could form the basis of a suite of different models, with population-specific nodes representing sensitive or otherwise unique populations represented as sub networks in the model.

4.5.4 Risk reduction

The fifth step of the risk management framework (NRC, 2009), risk management, has not been included in this modelling process, but is discussed here. In the context of wastewater reuse, risk management represents strategies to ameliorate the health risk of exposure to the treated wastewater. The WHO guidelines for wastewater irrigation recommend a change in focus from water quality standards at the point of release, to health-based targets at the point of exposure (WHO, 2006). These can be achieved step-wise, by a succession of risk reduction barriers that can be used in various combinations to reduce risk to an acceptable level befitting the end use of the water (Keraita et al., 2010). Under this broader, more flexible approach, the type and efficacy of wastewater treatment is not the sole barrier to infection. Therefore, in water recycling schemes where high levels of treatment are applied to reduce pathogen concentrations, lower levels of post-release controls are required to minimise pathogen concentrations and reduce exposure to hazards. Conversely, if lower levels of treatment are applied, then other barriers or methods to reduce exposure can be introduced (Cunliffe, 2006).

Post-treatment risk mitigation strategies for an agricultural irrigation scenario might be represented as a sub network, with nodes representing the steps or barriers to infection. These might include: restricting uses of recycled water (e.g., such as restricting use of water treated at a lower level to crops not eaten raw); controlling methods of application (e.g., underground vs. spray); setting withholding periods between application and harvesting (to allow pathogen die-off); or washing, cooking or peeling produce (NRMMC-EPHC-AHMC, 2006). In a similar fashion, although the existing model does not take into account the growth or inactivation of pathogens during storage or distribution of the treated wastewater, these elements could be modelled by the introduction of another sub network.

In summary, this conceptual model of the health risk of reusing wastewater from sewage maturation ponds has been developed from the relevant literature and using a participatory modelling process. In its present form, the model considers the health risk to an individual in a single exposure event, arising from exposure to the treated effluent at the point of discharge from a sewage maturation pond, incorporating the elements of the QMRA framework and the effects of the pond system on pathogen concentrations. Ultimately the aim of the model, once refined

and quantified, will be to use it for the purposes of prediction and risk mitigation. This work is an important addition to the formative work on assessing the public health impacts of recycled water use by Donald et al. (2009) and Cook et al. (2011), particularly due to its graphic elements, the close examination of QMRA and the consideration of post-treatment risk reduction measures. The model may be a useful 'drawing board' that can be further developed by a multidisciplinary stakeholder team comprising entities such as water utilities managers, regulatory authorities, hydraulic engineers and microbiologists, to assess health risk and evaluate risk reduction strategies pertaining to specific fit-for-purpose wastewater reuse scenarios.

Chapter 5: Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water

Preamble

This chapter has been written as a journal article to meet Objective 3 of the research, as stated in the Introduction:

Objective 3: To develop and evaluate a BN model for the assessment and management of microbial health risk in the context of wastewater reuse

Following the focused literature review described in Chapter 3 and the conceptual model detailed in Chapter 4, the next phase was to develop a prototype BN, based on a stochastic QMRA using data from peer-reviewed literature. Parameterisation of the network was achieved by computing conditional probabilities from QMRA model-generated data.

This chapter is primarily my own work and the figures and tables were created by me. The article was published by Science of the Total Environment in January, 2016 and is reproduced here in its entirety. The reference for the publication associated with this chapter is:

Beaudequin, D., Harden, F., Roiko, A., & Mengersen, K. (2016). Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water. Science of the Total Environment, 541, 1393–1409.

Contributor	Statement of contribution
D. Beaudequin	Conception of models, conduct of research, model development and population, data analysis, interpretation of results, writing of manuscript, modifications to manuscript as suggested by co-authors and reviewers
F. Harden	Comments on manuscript, editing
A. Roiko	Comments on manuscript, editing
K. Mengersen	Comments on manuscript, editing

Principal Supervisor Confirmation: I have sighted email or other correspondence for all co-authors confirming their authorship.

Name: KERRIE MENGERSEN Signature: _____ Date: 14/3/16

Abstract

Background: Quantitative microbial risk assessment (QMRA), the current method of choice for evaluating human health risks associated with disease-causing microorganisms, is often constrained by issues such as availability of required data and inability to incorporate the multitude of factors influencing risk. Bayesian networks (BNs), with their ability to handle data paucity, combine quantitative and qualitative information including expert opinions and ability to offer a systems approach to characterisation of complexity, are increasingly recognised as a powerful, flexible tool that overcomes these limitations.

Objectives: We present a QMRA expressed as a BN in a wastewater reuse context, with the objective of demonstrating the utility of the BN method in health risk assessments, particularly for evaluating a range of exposure and risk mitigation scenarios. As a case study, we examine the risk of norovirus infection associated with wastewater-irrigated lettuce.

Methods: A BN was developed following a QMRA approach, using published data and reviewed by domain experts using a participatory process.

Discussion: Employment of a BN facilitated rapid scenario evaluations, risk minimisation and predictive comparisons. The BN supported exploration of conditions required for optimal outcomes, as well as investigation of the effect on the reporting nodes of changes in 'upstream' conditions. A significant finding was the indication that if maximum post-treatment risk mitigation measures were implemented, there was a high probability (0.84) of a low risk of infection regardless of fluctuations in other variables, including norovirus concentration in treated wastewater.

Conclusion: BNs are useful in situations where insufficient empirical data exist to satisfy QMRA requirements and they are exceptionally suited to the integration of risk assessment and risk management in the QMRA context. They allow a comprehensive visual appraisal of major influences in exposure pathways and rapid interactive risk assessment in multifaceted water reuse scenarios.

5.1 INTRODUCTION

In many societies, the increasing demand that is being placed on water resources has motivated considerable interest and activity in water recycling. However there is a natural concern about the risk to human health associated with exposure to this type of water. Given the increasingly litigious and data-driven nature of many of these societies, the road to fit-for-purpose water reclamation and reuse continues to be labyrinthine. Policymakers err on the side of the precautionary principle and regulatory authorities are increasingly risk-averse, despite the irony that humanity is safer and healthier than at any time in history.

The assessment of human health risk from microbial and chemical exposure in wastewater reuse is multifaceted. Significant variation in the quality of the treated wastewater due to complex environmental influences, diverse and numerous exposure pathways depending on the reuse and variability in the dose response of human hosts, contribute to the conundrum that is the characterisation and quantification of risk in order to make informed decisions in the public health interest.

Bayesian networks (BNs), due to their many useful features, are emerging as an innovative technique in the assessment of microbial risk associated with foodborne and waterborne pathogens. A review of 15 publications by Beaudequin et al. (2015) examines the features of BNs that are particularly appealing in quantitative microbial risk assessment (QMRA), including their ability to incorporate diverse data types and to support ‘backwards’ or diagnostic reasoning (Beaudequin et al., 2015a). The small number of papers found was indicative of the novel use of the technique in QMRA. Furthermore, of the 15 papers reviewed, four were water-related and of these, only one studied recycled water, highlighting the nascent use of the method in evaluating risk and informing management options in the field of water reuse.

This paper describes the development and application of a probabilistic BN to characterise and quantify health risks associated with environmental exposures to pathogens in treated wastewater. The case study described in the paper focuses on a canonical exposure-risk pathway: the consumption-related health risk of norovirus infection associated with wastewater-irrigated lettuce. This provides the opportunity to showcase the BN approach, highlighting the potential of the methodology to augment traditional QMRA procedures. In addition to representing and quantifying

the comparative influence of key factors in the microbial exposure pathway, the BN features post-treatment risk reduction options and the capacity to simultaneously assess and quantify the effect of major factors thought to influence microbiological health risk.

We begin with a contextual description of the QMRA framework and a brief account of BN features. The paper includes an outline of the model construction and review process, followed by steps taken to validate the model and evaluate its sensitivity. Finally, a number of exposure scenarios are explored through the application of probabilistic inference and the determination of risk estimates for comparative purposes, thereby demonstrating the functionality of BNs and their attraction in microbial risk assessment and management.

5.1.1 Quantitative microbial risk assessment

QMRA is a framework for assessing public health risks from pathogenic organisms. The term describes a structured approach that uses data to describe the exposure and spread of microbial agents mathematically and quantify the nature of the adverse outcomes (CAMRA, 2013b, Havelaar, 2012, USEPA-USDA/FSIS, 2012). QMRA models generate knowledge about the propagation of microbiological hazards along the risk pathway from source to exposure in complex real-world scenarios. The primary purpose is to generate insight into the interdependence of input and output variables and to evaluate the effect of mitigation alternatives (Greiner et al., 2013). A QMRA approach is informative in the consideration of pathogenic risk because of the ability to examine many permutations of potential pathogens, sources of contamination and environmental influences that render epidemiological methods ineffective.

Although the QMRA concept and framework is widely recognised, there are acknowledged limitations (WHO, 2008). Microbial risk assessments are largely comprised of numerical simulation studies and are fundamentally data-dependent (Soller, 2012). These data are often limited, missing or poor in quality (Greiner et al., 2013, Petterson et al., 2001, World Health Organization, 2008), due to the microscopic nature of the subject. As a consequence, characterisation of exposure in complex real-world scenarios is difficult and an overall health risk assessment of a hazard considering all possible exposure routes is currently not possible due to data

gaps (Thoeye et al., 2003). In all, the many sources of uncertainty and variability are difficult to describe and separate (USEPA-USDA/FSIS, 2012).

5.1.2 Bayesian networks

BNs are a form of directed acyclic graph (DAG), where variables are represented by nodes and linked by arrows representing causality. The DAG provides an initial visual representation of relationships of influence among the variables. Each node or variable has a number of user-defined states that can be qualitative, discrete or continuous (e.g., 'true/false', 'high/low', '>5/≤ 5'). Each of the variable states is assigned a conditional probability, derived from empirical data (published or from field work), models, simulations, or expert opinion. The conditional probabilities reflect the strength of the causal links between the variables (Jensen and Nielsen, 2007) and further, the probability of each state in a 'child' node, that is, any node that has a node directly above it, is conditional upon the influences of its 'parent' nodes. Software supporting BNs present node states and their conditional probabilities in a table underlying each node. The network efficiently expresses the joint probability distribution of the variables and changes in the network as a result of adding new information are immediately reflected in the response node, or outcome of interest (Greiner et al., 2013).

BNs are one of a number of approaches to 'integrated' or transdisciplinary modelling in environmental applications. Other methods using a systems approach include meta models, coupled complex models, system dynamics, agent based models and expert systems (Jakeman et al., 2007). In common with other methods, BNs offer visualisation as network graphs and the expression of parameter uncertainty using distribution functions. Importantly however, BNs have an added advantage of conditional independence of the nodes, in that the probability distribution for any one node depends only on its parent nodes, therefore simplifying quantification and expediting inferential reasoning.

Due to these features BNs are able to characterise a complex issue, quantify outcomes of interest and describe the many possible interactions between variables (Donald et al., 2009). The method can be used for 'forward inference', where inputs are specified and impact on outcome is observed, thus revealing variables that are major drivers for an outcome. Through Bayes' theorem (Jensen and Nielsen, 2007, Pearl, 2000), BNs are also capable of diagnostic or 'backwards reasoning', by which

the outcome is specified and the states of the system's variables required to obtain that outcome are calculated (Ben-Gal, 2007, Coupé et al., 2000, Pollino and Hart, 2005). As each variable in a BN is represented as a probability distribution, uncertainty is made explicit. This attribute is particularly valuable when modelling environmental systems, as uncertainty can be widespread (Aguilera et al., 2011). These attributes, combined with the ability to combine different data types and expert opinion, allow the use of BNs to address key problem areas in QMRA such as difficulties with modelling complex exposure pathways, lack of or poor data in dose-response modelling and the necessity to characterise uncertainty to produce a credible risk estimate (Beaudequin et al., 2015a). Importantly, a QMRA can provide inputs into a BN or a BN can be used to augment a QMRA (Donald et al., 2009), including expressing an entire stochastic QMRA model as a BN (Greiner et al., 2013, Rigaux et al., 2012a, Smid et al., 2010).

5.1.3 The pathogen

For this case study norovirus was selected as a 'worst case' pathogen, since it is highly infectious (Ong, 2013), persistent in the environment (Ong, 2013, Silverman et al., 2013) and resistant to wastewater treatment (Da Silva et al., 2008, NRMMC-EPHC-AHMC, 2006, Symonds et al., 2014). Moreover, noroviruses infect people of all ages (Widdowson et al., 2005). Worldwide, noroviruses cause over 90% of all viral gastroenteritis cases (Patel et al., 2009) and in Australia they are the most common cause of gastroenteritis, causing an estimated 1.8 million cases per annum (Hall et al., 2005).

5.1.4 The exposure scenario

Salad crops, particularly lettuce, feature commonly in case studies of exposure to waterborne pathogens associated with wastewater-irrigated food crops, representing a conservative scenario as they are generally eaten raw. The relationship between norovirus dose and adverse human health effects has been clearly identified as a knowledge gap (USFDA, 2013), highlighting the importance of investigating aspects of this association to contribute to the knowledge repository on this pathogen (Ong, 2013). The scenario of norovirus contamination on wastewater-irrigated lettuce has been widely studied (Bae et al., 2011, Barker, 2014a, Barker et al., 2014, Barker et al., 2013a, Mara and Sleigh, 2010, Mok et al., 2014, Sales-Ortells et al.,

2015), offering an abundance of data with which to parameterise, validate and evaluate a BN examining influences on norovirus infection risk.

In the following paragraphs, section 5.2.1 of the Method describes the assumptions of the case study; section 5.2.2 outlines the underlying QMRA process models and sections 5.2.3 and 5.2.4 describe the construction and validation of the BN.

5.2 METHOD

5.2.1 QMRA

A deterministic model of the risk of norovirus infection and illness was constructed first in Microsoft Excel and validated using point estimates from the literature. A stochastic model was then developed with distributions derived from peer-reviewed studies, or triangular distributions using maxima and minima from the literature where more specific distributions were not available. Model input parameters can be found in Table C1 in Appendix C. A dataset for each variable was generated by MC simulation with random sampling of 10,000 values from each distribution input using @Risk 6.2 (Palisade Corporation).

Modelling norovirus concentrations in treated wastewater

Norovirus is particularly difficult to enumerate (Atmar, 2010), due to difficulties with detection and recovery, high assay costs and current incapacity to culture the virus. For this case study, the estimated norovirus concentration after treatment was derived from a range of norovirus concentrations in raw sewage of 10^4 to 10^7 PCR units per mL reported by Barker (2014a), to which the 6 log unit reduction recommended for viral pathogens during wastewater treatment (NRMMC-EPHC-AHMC, 2006) was applied as follows:

$$C_{treat} = 10^{(\log_{10} C_{raw} - LRV)} \quad (1),$$

where C_{treat} is the concentration of norovirus in the treated wastewater (PCR units/mL), C_{raw} is the concentration of norovirus in raw sewage and LRV is the $\log_{10}$ unit reduction of the pathogen achieved through treatment. The mean of the resulting range was used to model the norovirus density in treated wastewater as a lognormal distribution (Tanaka et al., 1998, Westrell et al., 2006).

Modelling the exposure pathway

In this case study, it was assumed that treated wastewater alone, delivered via sprinkler system every 3-6 days is used for irrigation of an established lettuce crop (DAF QLD, 2010). Viral accumulation resulting from successive irrigation events was assumed to be offset by viral decay (Mok et al., 2014). The exposure route is direct ingestion through lettuce consumption. Ingestion is considered to be the route carrying the highest risk for recycled water, since it delivers the largest dose of pathogens (EPA QLD, 2005). This assumption may vary however, depending on the scenario and or the pathogen (Thoeye et al., 2003); for example, in the case of a farm worker exposed to spray irrigation with recycled water, the most risk-laden route may be inhalation or inhalation-ingestion.

Modelling risk mitigation measures

In keeping with recommendations by WHO guidelines for wastewater use in agriculture (2006) for a change in emphasis from water quality at the point of release to actual pathogen dose at the point of exposure, this study incorporated two post-release risk reduction measures (Keraita et al., 2010). These were pathogen reduction through natural die-off as a result of withholding irrigation prior to harvest (O'Toole, 2011, Petterson et al., 2001, Petterson et al., 2002) and lettuce washing (Bae et al., 2011, Dawson et al., 2005, Predmore and Li, 2011). Using a BN, the influence and interaction of risk mitigating factors in a number of scenarios can be examined rapidly. These might include simulation of water recycling schemes where low levels of water treatment are applied to reduce pathogen concentrations, in conjunction with higher levels of post-release controls to reduce exposure to hazards (Cunliffe, 2006).

5.2.2 QMRA process models

The unmitigated dose D_u (PCR units) of norovirus resulting from the consumption of lettuce irrigated with treated wastewater was modelled as:

$$D_u = VLC_{treat} \qquad (2),$$

where V is the volume of treated wastewater retained on lettuce leaves (mL/g), L is the mass of lettuce (g) consumed in one exposure event and C_{treat} is the concentration of norovirus in the treated wastewater (PCR units/mL). The pathogen reduction effect of ceasing irrigation for a designated number of days prior to

harvesting the lettuce to allow pathogen die-off to occur) $R_{withholding}$ was modelled as:

$$R_{withholding} = e^{(-k_{decay}\, t)} \quad (3)$$

(Mok et al., 2014), where k_{decay} is the pathogen decay constant per day and t is the withholding period (days), i.e., the number of days between the last irrigation of lettuce with treated wastewater and harvesting. The pathogen reduction effect produced by washing lettuce prior to eating ($R_{washing}$) is given by:

$$R_{washing} = 10^{(-LRV_{washing})} \quad (4)$$

(WRA, 2014a), where $LRV_{washing}$ is the $\log_{10}$ reduction in pathogen load due to lettuce washing. The total pathogen reduction R_{total} is given by:

$$R_{total} = R_{withholding} R_{washing} \quad (5)$$

(Barker, 2014a). The mitigated norovirus dose D_m (PCR units) is given by:

$$D_m = R_{total} D_u \quad (6)$$

(Barker, 2014a), where R_{total} is the combined pathogen reduction effect of lettuce washing and the withholding period and D_u is the unmitigated dose of norovirus. The risk of infection per exposure event, P_{inf}, was modelled as:

$$P_{inf} = 1 - \left(1 + \frac{D_m}{N_{50}}\left(2^{\frac{1}{\alpha}} - 1\right)\right)^{-\alpha} \quad (7)$$

(Teunis et al., 2008), where D_m is the mitigated dose, N_{50} is the median effective dose and α is the slope parameter. The values used for N_{50} and α were 16963 and 0.110891, respectively (Teunis et al., 2008). The annual probability of infection, $P_{\mathrm{inf}(ann)}$ is given by

$$P_{\mathrm{inf}(ann)} = 1 - \left(1 - P_{inf}\right)^{F} \quad (8),$$

where F denotes the annual frequency of lettuce consumption in days per year. The annual probability of illness $P_{ill(ann)}$ is given by

$$P_{ill(ann)} = P_{inf(ann)} R_{ill:inf} \quad (9),$$

where $R_{ill:inf}$ denotes the illness to infection ratio.

5.2.3 Bayesian network

The structure of the BN model was derived from the QMRA process models described above, using GeNie 2.0 (UP DSL, 2013). Threshold values and ranges for states (Table C2 in Appendix C) were selected on the basis of indicative or typical values for Australian conditions. Other options for node discretisation could include using equal bin size or equal probabilities. The states of parentless or root nodes were parameterised using prior knowledge of the likely frequencies of each state based on Australian conditions from the literature, or uniform probabilities if the prior distribution was unknown. Parameterisation of the states of child nodes was accomplished by calculating conditional probabilities from the simulated data set generated from 10,000 iterations of the stochastic process models.

The *Risk of infection* node (Figure 5.1) provides single-event information on the probability of infection for an individual eating lettuce, where infection is defined as faecal excretion of virus and seroconversion (Teunis et al., 2008). The *Annual risk of infection* node gives an annualised estimate of infection risk for an individual eating lettuce over a period of one year and correspondingly, the *Annual risk of illness* node provides information on the risk of illness in an individual eating lettuce over a period of one year. Illness is defined as the development of symptoms of diarrhoea and/or vomiting combined with other symptoms (Teunis et al., 2008). These three nodes are the response nodes in the BN.

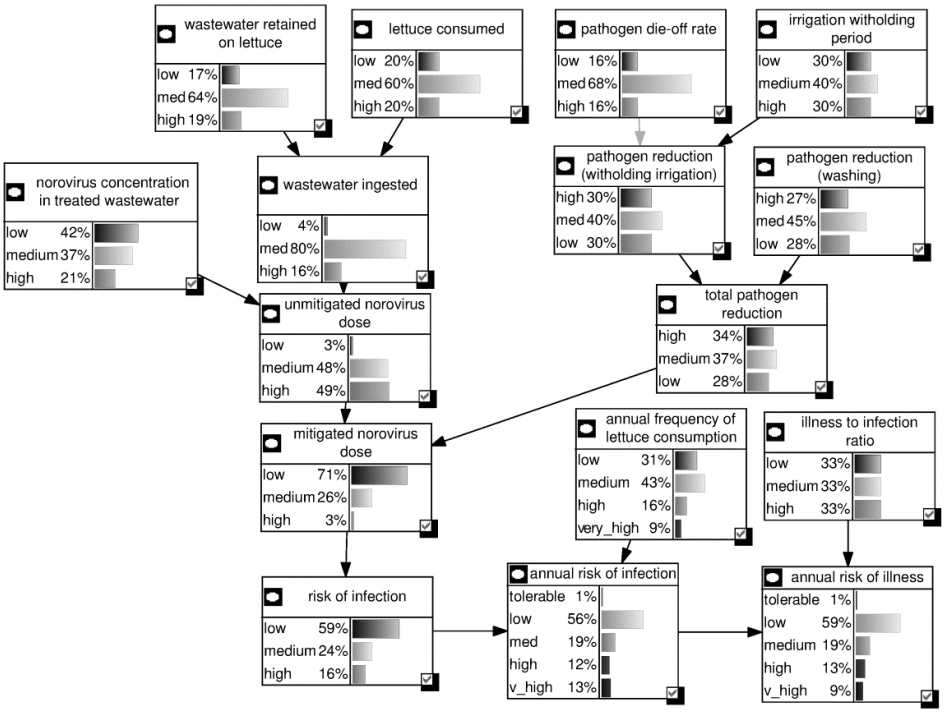

Figure 5.1. Bayesian network of risk of norovirus infection and illness from consumption of wastewater-irrigated lettuce.

The network in Figure 5.1 presents the prior state of knowledge of the risk of norovirus infection from eating lettuce irrigated with treated wastewater. As discussed above, the prior probability distribution of each child node was computed from the relationships of influence and resulting conditional probabilities in the underlying stochastic model. Figure 5.1 thus represents a 'baseline' status of the network, i.e., before any further specific evidence is added. At the baseline level of knowledge, the value assigned to a node state represents the chance that the node will be in that state. For example, in the *Wastewater ingested* node, the chance that a 'Medium' level of wastewater will be ingested each time a serving of lettuce is consumed is 80%; in the *Risk of infection* node, the baseline values show that the chance of having a low risk of infection is 59%, thus being the most likely outcome. This nonetheless means a 41% chance of a medium or high infection risk. The annual risks of infection and illness have a correspondingly low probability (0.01 each) of being at or below the established 'tolerable' threshold. The tolerable annual risk of infection and illness for norovirus have been estimated to be 1.4×10^{-3} and 1.1×10^{-3} per person per year (pppy) respectively (Mara and Sleigh, 2010), based on a tolerable DALY loss of 10^{-6} pppy for additional burden of disease as a result of wastewater

use in agriculture (WHO, 2006). The debate about the appropriateness of these established tolerable risk thresholds is expounded in the Discussion.

5.2.4 Validation of the BN

Validation of the BN followed the principles outlined by Pitchforth and Mengersen (2013) and Pollino et al. (2007). In the structural development and evaluation phase, an unparameterised causal network was constructed initially from the peer-reviewed literature (Beaudequin et al., 2015b) and reviewed by representatives from the water industry and state health authority and researchers from the disciplines of microbiology and environmental health. Based on this conceptual model, a refined BN comprising key variables was developed, quantified and validated by a subgroup of the same stakeholders. Quantitative validation of the deterministic QMRA process models was undertaken during this stage, by comparing model output with published likely values. The BN model structure and discretisation and parameterisation of nodes were also validated by specialist participants who were not involved in creating the model (Pitchforth and Mengersen, 2013). The refined model was presented to and critically evaluated by an academic audience. Validation of model behaviour was accomplished by cross-checking the probabilities in the network with each other for consistency, such that if the probabilities in one node of the network were changed, the subsequent changes in probabilities of other nodes followed approximately as expected (Pitchforth and Mengersen, 2013). Sensitivity analysis, described in the next section, completed the model validation procedures, in the absence of other BNs in this domain with which to compare model output. Sensitivity to findings, including that model endpoints were more sensitive to lettuce washing than wastewater retained on leaves, reflected QMRA model findings reported by Barker (2014a).

One of the attributes of a BN is the ability to determine which nodes are both modifiable and most influential in achieving the desired response or outcome of interest. In the ensuing paragraphs we simulate a number of complex, hypothetical exposure scenarios and their effect on the outcomes of interest, thereby demonstrating the utility of the BN in the rapid assessment of wastewater reuse (or water use) contexts.

5.3 RESULTS

5.3.1 Sensitivity assessment of the BN

A preliminary visual diagnostic procedure showed that the *Irrigation withholding period* and *Pathogen reduction (washing)* nodes had the most influence on the target node *Risk of infection,* followed by the *Norovirus concentration in treated wastewater* node.

Sensitivity to findings was assessed by determining those root nodes that most strongly influence a selected target variable. Root nodes were set independently to the maximum optimal conditions. The sensitivity analysis showed that withholding irrigation and lettuce washing had the most influence on risk of infection, followed by wastewater pathogen concentration. The amount of lettuce consumed and volume of wastewater retained on lettuce leaves had modest effects on infection risk whereas differences in pathogen decay rates had no effect on infection risk. Table 5.1 shows the influence of root variables on the *Risk of infection* node in rank order from the variables with the largest influence on the target node. The prior probabilities referred to in the Table 5.1 caption were derived from Figure 5.1.

The following scenarios are included to demonstrate the versatility of BN models for rapid evaluation of potential events influencing water reuse decisions. The scenarios also serve to illustrate the functional attributes of the BN; Scenario 1 uses ‘backward inferencing’, Scenario 2 demonstrates the utility of simultaneous changes in multiple nodes, Scenarios 3-5 use ‘forward inferencing’ and Scenario 6 illustrates both changes in multiple nodes and the capacity for sequential change assessment.

Table 5.1

Sensitivity to Bayesian network findings for root variables in rank order from variables with most influence (heavily shaded cells) to variables with least influence (unshaded cells). Prior probability for low Risk of infection was 0.59 and high was 0.16

root node	risk of infection posterior probability		risk of infection difference in probability	
	low	high	low	high
Irrigation withholding period: 100% high	0.83	0.01	0.24	-0.15
Irrigation withholding period: 100% low	0.33	0.34	-0.26	0.18
Pathogen reduction (washing): 100% high	0.77	0.04	0.18	-0.12
Pathogen reduction (washing): 100% low	0.44	0.27	-0.15	0.11
Norovirus concentration in treated wastewater: 100% high	0.48	0.24	-0.11	0.08
Norovirus concentration in treated wastewater: 100% low	0.68	0.10	0.09	-0.06
Lettuce consumed: 100% high	0.55	0.19	-0.04	0.03
Lettuce consumed: 100% low	0.62	0.14	0.03	-0.02
Wastewater retained on lettuce: 100% high	0.57	0.18	-0.02	0.02
Wastewater retained on lettuce: 100% low	0.62	0.14	0.03	-0.02
Pathogen die-off rate: 100% high	0.59	0.16	0	0
Pathogen die-off rate: 100% low	0.59	0.16	0	0

5.3.2 Scenario assessments

Scenario 'Tolerable annual risk'

In the first instance it is of interest to determine the network conditions under which it is certain that a tolerable annual risk of infection will not be exceeded. This goal is simulated by setting *Annual risk of infection* to 100% tolerable (Figure 5.2). The subsequent changes in states required in modifiable nodes to achieve this outcome are shown in Table 5.2. The modifiable variable with the largest influence in achieving the target of 100% certainty for a tolerable annual risk of infection is the

Irrigation withholding period node. The BN indicates that exposure factors such as the volume of wastewater retained on lettuce leaves and the size of the lettuce serving (mass) were not as influential in achieving an overall low risk of infection, as the pathogen concentration in the reclaimed water and the total pathogen reduction through post-treatment barriers such as lettuce washing and withholding irrigation. This scenario is an example of the 'backwards reasoning' ability of BNs, whereby although the *Risk of infection* node was quantified by determining its conditional probabilities given the states of its parent nodes. The use of priors and Bayes' theorem allows the probability of the states of the input nodes to be determined given a defined outcome.

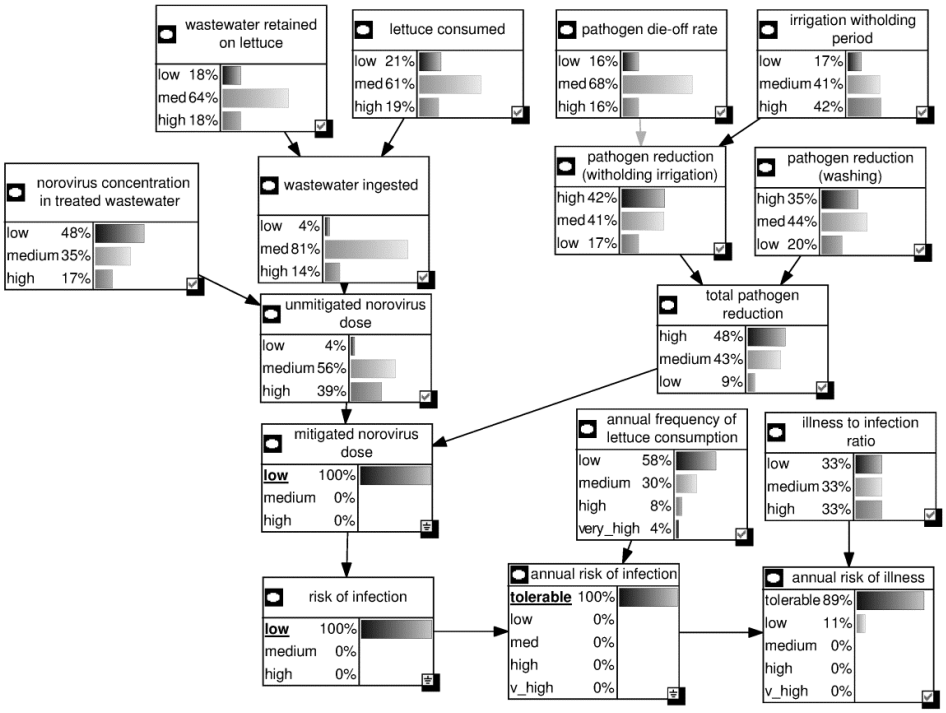

Figure 5.2. Bayesian network for Scenario 'Tolerable annual risk', displaying variable conditions required for certainty of a tolerable annual risk of infection.

Table 5.2

Scenario 'Tolerable annual risk', displaying changes required in Bayesian network modifiable nodes for certainty of a tolerable annual risk of infection. Pr(medium) = 1 – Pr(low or high)

	baseline probability		probability required		difference	
modifiable node	**low**	**high**	**low**	**high**	**low**	**high**
Irrigation withholding period	0.30	0.30	0.17	0.42	-0.13	0.12
Pathogen reduction (washing)	0.28	0.27	0.20	0.35	-0.8	0.08
Norovirus concentration in treated wastewater	0.42	0.21	0.48	0.17	0.06	-0.04
Wastewater retained on lettuce	0.17	0.19	0.18	0.18	0.01	-0.01

Scenario 'Outbreak'

In this scenario a small community serviced by a sewage maturation pond experiences an outbreak of norovirus infection, resulting in raised concentrations of norovirus in the treated wastewater, sold locally to irrigate a commercial lettuce crop. This evidence was added to the network by setting the *Norovirus concentration in treated wastewater* node to 100% 'high'. Although the chance of a 'high' risk of infection increased substantially by 50%, the probability of achieving a tolerable annual risk of infection or illness remained unchanged at 0.01. Two other response nodes also underwent substantive changes. These were *Annual risk of infection* and *Annual risk of illness*, which showed 39% and 33% increase respectively in the chance of being 'very high'. There were no changes in the other root nodes in response to the introduced evidence in *Norovirus concentration in treated wastewater*, due to the property of 'd-separation' (Pearl, 1988), wherein given the conditions in the evidentiary node *Norovirus concentration in treated wastewater*, these nodes add no further information to what is known about the response nodes. The changes in response nodes are illustrated in Figure 5.3. The full BN for this scenario can be found in Figure D1 in Appendix D.

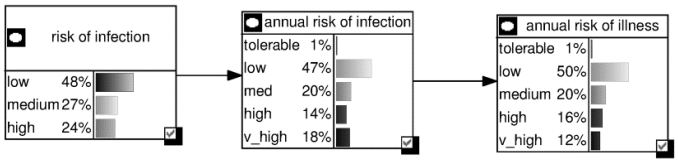

Figure 5.3. Scenario 'Outbreak', displaying response nodes for risk of norovirus infection.

Scenario 'Outbreak with risk mitigation'

In response to the outbreak, the water utility authority advises farmers to withhold irrigation for three days before harvesting for market, to allow natural pathogen die-off to occur. This evidence is introduced to the network by setting the *Irrigation withholding period* node to 100% 'high'. The BN model indicates this measure is very effective, as the chance of a 'high' infection risk drops by 96% from the level in the outbreak scenario. The chance of a 'tolerable' annual risk of both infection and illness increased by 100% and that of a 'very high' annual risk of infection and illness were both reduced by 83%. While there was logically no change in the *Unmitigated norovirus dose* node in response to the risk reduction measure, the *Mitigated norovirus dose* node showed a 67% increase in the chance of a 'low' dose and 92% and 100% reductions in the chance of a medium and high dose, respectively, from the levels in the outbreak scenario. The changes in response nodes are illustrated in Figure 5.4. The full BN for this scenario can be found in Figure D2 in Appendix D.

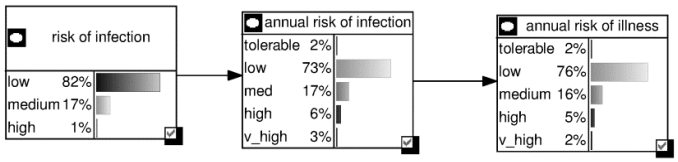

Figure 5.4. Scenario 'Outbreak with risk mitigation', displaying response nodes for risk of norovirus infection.

Scenario 'Furrow system'

A lettuce farmer wants to evaluate the worth of changing from spray irrigation to a furrow system that reduces the volume of wastewater captured by the leaves of the crop. This scenario is simulated by setting the *Wastewater retained on lettuce* node to 100% 'low', resulting in a 300% increase in the likelihood of the wastewater ingested by a lettuce consumer being low and causing a 133% increase in the chance of the unmitigated norovirus dose being low. However since the response nodes were not exceptionally sensitive to the *Wastewater retained on lettuce* node (as seen in

Table 5.1), the most significant changes in the response nodes were a reduction in the chance of a 'high' risk of infection by 12% and a corresponding reduction in the chance of a 'very high' annual risk of infection and illness of 15% and 11%, respectively. In this case the simulated outcome was judged by the farmer to be insufficiently different to warrant the cost of changing the irrigation system. The changes in response nodes are illustrated in Figure 5.5. The full BN for this scenario can be found in Figure D3 in Appendix D.

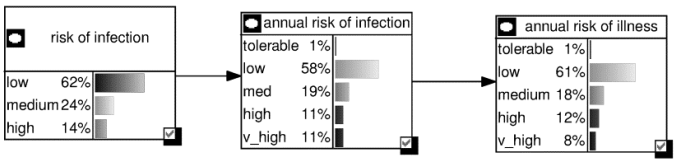

Figure 5.5. Scenario 'Furrow system', displaying response nodes for risk of norovirus infection.

Scenario 'Treatment change'

After the norovirus epidemic in the 'Outbreak' scenario has resolved, the water utility authority considers adding a reed bed filtration step to the existing maturation pond system to enhance its treatment capabilities (Jackson and Jackson, 2008), reducing the norovirus concentration after treatment from the baseline value. The efficacy of this approach is modelled by setting the *Norovirus concentration in treated wastewater* node to 100% 'low'. A significant difference was seen in the chance of a 'high' risk of infection, which decreased by 38%, with a corresponding reduction in the chance of a 'very high' annual risk of infection and illness by 38% and 33%, respectively. The chance of a 'low' unmitigated norovirus dose increased by 133% and the chances of a 'high' unmitigated norovirus dose was reduced by 73%. The most significant effect on the mitigated norovirus dose was a 67% reduction in the chance of a 'high' dose. The changes in response nodes are illustrated in Figure 5.6. The full BN for this scenario can be found in Figure D4 in Appendix D.

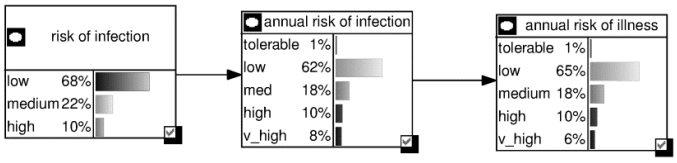

Figure 5.6. Scenario 'Treatment change', displaying response nodes for risk of norovirus infection.

Scenario 'Lettuce washing'

A regulation change leads to the requirement that lettuces irrigated with treated wastewater are washed before sending to market. As a result of adding this evidence to the network by setting the *Pathogen reduction (washing)* node to 100% 'high', the chance of a 'high' infection risk is reduced by 75%, with a corresponding doubling of the 'tolerable' annual risk of infection. Although there was no change in the probability of achieving a 'tolerable' annual risk of illness, there was a 62% reduction in the chance of a very high annual risk of infection and a 67% reduction in the chance of a very high annual risk of illness. As expected the nodes affected by the new evidence were the *Mitigated norovirus dose* node and the *Total pathogen reduction* node. The former displayed a 100% decrease in chance of a high dose and a 65% reduction in chance of a medium dose, while the latter node showed a 100% decrease in the chance of low pathogen reduction and a corresponding increase in chance of high and medium pathogen reduction of 56% and 27% respectively. The changes in response nodes are illustrated in Figure 5.7. The full BN for this scenario can be found in Figure D5 in Appendix D.

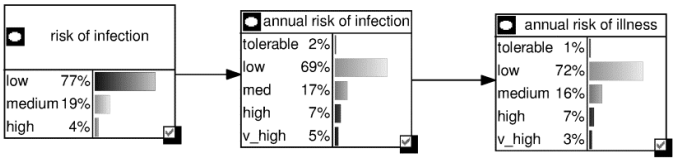

Figure 5.7. Scenario 'Lettuce washing', displaying response nodes for risk of norovirus infection.

Scenario 'Rain'

A severe rain event leads to more pathogens being washed from the catchment into the sewage treatment plant, but this also results in less reliance on recycled water for irrigation. This evidence is added to the network by setting *Norovirus concentration in treated wastewater* node to 100% 'high' and *Wastewater retained on lettuce* node to 100% 'low'. Despite reduced likelihood of wastewater on lettuce, chance of a 'high' infection risk increased by 44%, with corresponding increases in the chance of a 'very high' annual risk of infection and illness of 31% and 33% respectively. Notwithstanding the chance of 'low' ingestion of wastewater increasing by 300% and chance of high wastewater ingestion reducing by 94%, there was still a 100% reduction in chance of a 'low' unmitigated norovirus dose and a 100% increase in chance of a 'high' mitigated norovirus dose even with reduced probability

of wastewater ingested. This apparent contradiction reflects relatively smaller strength of influence exerted by the *Wastewater retained on lettuce* node on response nodes. The changes in response nodes are illustrated in Figure 5.8. The full BN for this scenario can be found in Figure D6 in Appendix D.

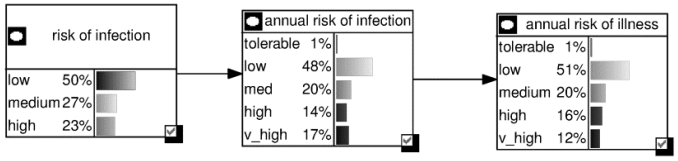

Figure 5.8. Scenario 'Rain', displaying response nodes for risk of norovirus infection.

Scenario 'Rain with decreased withholding period'

In the absence of any information to the contrary, the lettuce farmer reasons that because there has been very little irrigation due to rain and subsequent risk of pathogen contamination is low, weather-related crop losses can be compensated for by reducing the withholding period and harvesting remaining lettuce immediately. This evidence is added to the evidence from the 'Rain' scenario by setting *Irrigation withholding period* node to 100% 'Low', resulting in the chance of a 'high' risk of infection increasing by 100%. A 'tolerable' annual risk of infection and illness is now also unachievable, as the chance of the tolerable state in both cases has dropped to 0%. The changes in response nodes are illustrated in Figure 5.9. The full BN for this scenario can be found in Figure D7 in Appendix D.

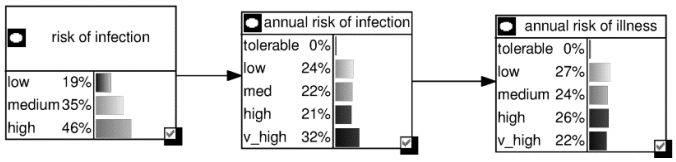

Figure 5.9. Scenario 'Rain with decreased withholding period', displaying response nodes for risk of norovirus infection.

The chances of achieving each of the response node states for the seven scenarios and baseline conditions described above are summarised in Table 5.3. 'Outbreak with risk mitigation' was the best case scenario, providing the highest chance (82%) of a low risk of infection from a single exposure event. This result was reflected in the alternative outcome measure annual risk of illness, in which 'Outbreak with risk mitigation' also resulted in the highest chance (2%) of a tolerable annual risk of illness. Of the scenarios considered, 'Rain with decreased withholding period' represented the worst case, with a 46% chance of a high risk of infection

from a single exposure event and a 32% chance of a very high annual risk of infection. A tolerable annual risk of infection and illness was not achievable in this scenario.

Table 5.3

Chance of achieving each response node state for seven scenarios and baseline conditions

scenario	risk of infection (%)			annual risk of infection (%)					annual risk of illness (%)				
	low	medium	high	tolerable	low	medium	high	very high	tolerable	low	medium	high	very high
Baseline	59	24	16	1	56	19	12	13	1	59	19	13	9
Outbreak	48	27	24	1	47	20	14	18	1	50	20	16	12
Outbreak with risk mitigation	82	17	1	2	73	17	6	3	2	76	16	5	2
Furrow system	62	24	14	1	58	19	11	11	1	61	18	12	8
Treatment change	68	22	10	1	62	18	10	8	1	65	18	10	6
Lettuce washing	77	19	4	2	69	17	7	5	1	72	16	7	3
Rain	50	27	23	1	48	20	14	17	1	51	20	16	12
Rain with decreased withholding period	19	35	46	0	24	22	21	32	0	27	24	26	22

In addition to the analyses presented in the scenarios above, a significant finding in implementation of this BN was that if post treatment risk reduction was set to maximum effect i.e., if *Irrigation withholding period* and *Pathogen reduction (washing)* nodes were set to 100% 'high', the chance of a low risk of infection increased to 84% (Figure 5.10). This chance remained unchanged regardless of all other variables, alone or in combination, being modelled as high or low. In other words, the BN indicates that if maximum post-treatment risk mitigation measures were implemented, chance of a low risk of infection would always be 84%, regardless of any changes to other variables, including norovirus concentration in treated wastewater.

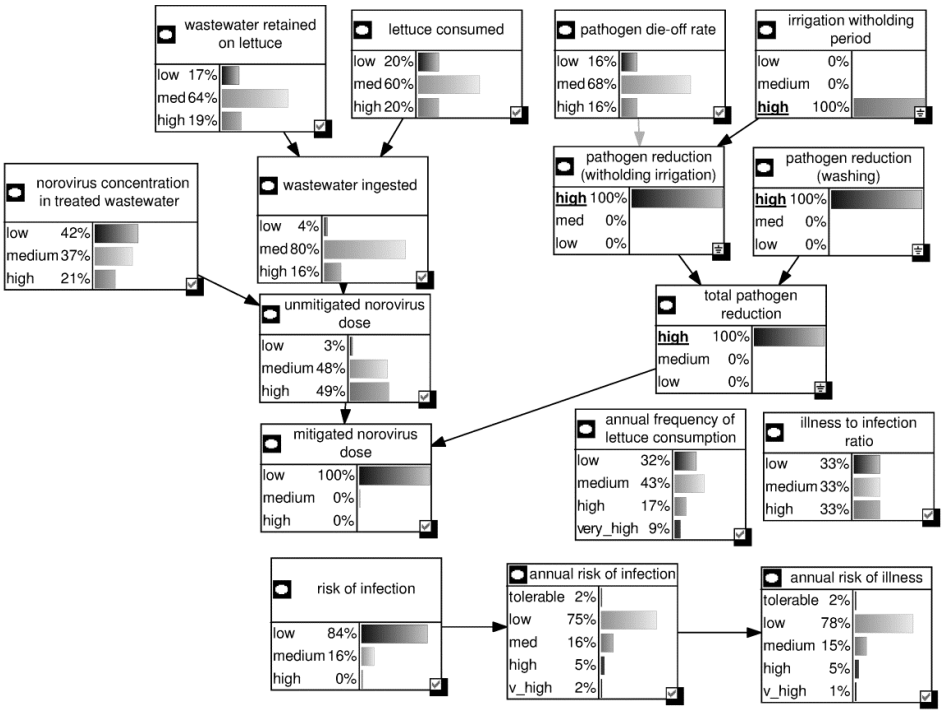

Figure 5.10. Bayesian network simulating maximum post-treatment risk reduction measures.

As indicated previously, threshold values for the 'Tolerable' states in *Annual risk of infection* and *Annual risk of illness* nodes (1.4×10^{-3} and 1.1×10^{-3} pppy respectively) estimated by Mara and Sleigh (2010), are based on a tolerable DALY loss of 10^{-6} pppy (WHO, 2006). In its prior state, the BN (Figure 5.1) indicates 1% chance of achieving a tolerable annual risk of norovirus infection and the most likely outcome, with a probability of 0.56, is a low risk of infection, defined as > 0.0014 and ≤ 0.2510 pppy and exceeding the tolerable threshold. However Mara (2011) argues compellingly from a number of perspectives that the tolerable DALY loss on which these estimates are based is too stringent (Mara, 2011). He puts a strong case for consideration of a tolerable DALY loss of 10^{-4} pppy, which would alter the norovirus tolerable annual risks of infection and illness to 1.4×10^{-1} and 1.1×10^{-1} pppy, respectively.

Figure 5.11 displays a BN modified to reflect the change in tolerable DALY loss proposed by Mara (2011). In its prior state this modified BN now expresses a probability of 0.45 that tolerable annual infection risk for norovirus will not be exceeded, now the most likely outcome. Correspondingly, the probability for a tolerable annual risk of illness is 0.42. These modifications demonstrate the sensitivity of the response nodes to the threshold values of their states, evidenced by

the dramatic changes in probabilities when threshold values are varied from the established tolerable risk level (Mara and Sleigh, 2010) to the alternative proposed by Mara (2011).

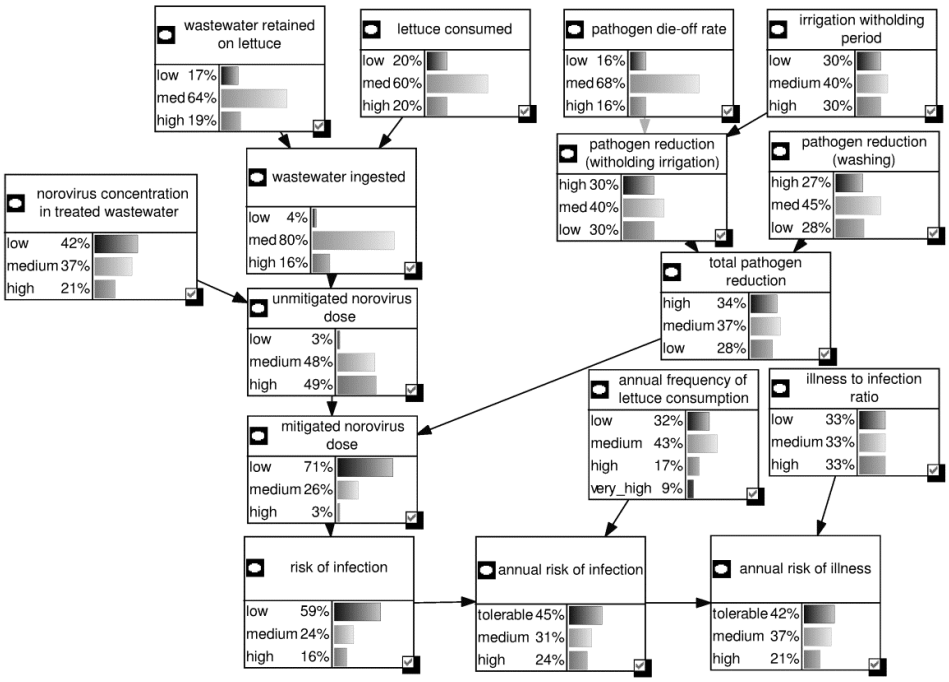

Figure 5.11. Modified BN with tolerable risk thresholds for *Annual risk of infection* and *Annual risk of illness* nodes reflecting change in tolerable DALY loss from 10^{-6} to 10^{-4} proposed by Mara (2011).

5.4 DISCUSSION

BNs and other QMRA models can be constructed and parameterised in a number of ways to encompass the comparative nature of risk assessment, i.e., an ability to compare relative risks. Moreover, the ability of BNs to present and evaluate multiple models and parameterisations is particularly appealing for better understanding of complex systems. QMRA is acknowledged to be most informative when used to compare relative risks of scenarios or decision options, due to the significant variability and uncertainty associated with assessing absolute risk from microorganism exposure (Soller et al., 2004). As we have shown, this optimal use of the QMRA method is rapidly achieved with a BN, which has the additional benefit of instantly updating when new information is provided and presenting outcomes, influences and options in a straightforward manner to users from any discipline via a convenient visual platform.

In previous applications of the approach, a BN was used as a supplementary analysis to a QMRA to identify the nodes with the most influence on incidence of gastroenteritis associated with recycled water (Donald et al., 2009). Since this formative work, four studies have subsequently used BNs in the assessment of water-related microbial risk, although none were on recycled water. Researchers aiming to reduce uncertainty in management decisions regarding potential threat of faecal contamination in recreational waters used a BN to explore differences between sampling locations and times and between analytical methods for quantifying FIB concentrations (Gronewold et al., 2011). A study of small private water supplies in England and France (Hunter et al., 2011), incorporated a pathogen-*E.coli* regression model in the BN, revealing very high infection risks to consumers from *Cryptosporidium* and *Giardia*. A BN was used in a QMRA of waterborne pathogens in a freshwater lake to predict levels of human health risk from factors such as physicochemical parameters and faecal indicator bacteria densities (Staley et al., 2012) and a BN was used with QMRA to prioritise public health management options for wet weather sewer overflows (Goulding et al., 2012). The full potential of the BN approach, with benefits for multidisciplinary teams and the exploration of complex systems, has yet to be realised.

As indicated earlier, risk of norovirus infection associated with consumption of leafy vegetables has been studied extensively by others. A recent study assessing the risk of norovirus infection from eating lettuce and other vegetables, irrigated with treated wastewater, found that vegetable washing significantly reduced risk (Barker, 2014a). This finding was reflected in the BN outcome in the 'Lettuce washing' scenario, wherein the chance of a tolerable annual risk of infection doubled with lettuce washing. The same study also demonstrated significant risk reduction with a one-day irrigation withholding period. The BN *Risk of infection* node correspondingly showed a 6% reduction in the chance of a high risk of infection resulting from a one-day irrigation withholding period, however more noticeably, a 94% reduction in the chance of a high risk of infection was achieved with a three-day withholding period, translating to doubling of the tolerable annual risk of infection and illness. Sensitivity analysis of the QMRA in the study revealed that uncertainty predominated in factors such as pathogen log removal during treatment, rates of

vegetable consumption and pathogen reduction due to vegetable washing (Barker, 2014a).

The QMRA model underlying the BN, assuming universal lettuce washing, yielded a median risk of infection per dose of 1.34×10^{-3}, which was comparable to the upper bound of the range for vegetable washers in another paper on norovirus disease burden from wastewater irrigation of vegetables (Mok et al., 2014). Uncertainty in the virus kinetic decay constant in the latter study did not significantly contribute to variation and this finding was reflected in the low sensitivity to pathogen die-off rate in the BN. A similar investigation of health risks associated with consumption of lettuce irrigated with treated effluent (Sales-Ortells et al., 2015) revealed prevailing model sensitivities to consumption of lettuce and concentration of norovirus in the treated effluent. In contrast, the BN found irrigation withholding period and lettuce washing were the most influential factors on measures of infection risk.

QMRA has a pathogen-specific approach to risk assessment, predicting disease outcomes prospectively as a result of exposure to a single pathogen, whereas contact with wastewater potentially entails simultaneous exposure to multiple pathogens. This limitation can be overcome partially by conducting a QMRA for reference pathogens from each of the three key pathogen groups – bacteria, viruses and protozoa (Mara, 2011). One of the attractions of BNs is their quantification using data in different formats from diverse sources, including opinions obtained from those with specialist knowledge. Inclusion of expert knowledge or epidemiological or qualitative data in a BN can assist in the evaluation of more generalised model endpoints such as overall morbidity from gastroenteritis, thus accounting for exposure to multiple pathogens. Use of BNs in risk assessments of water quality need not, therefore, be limited to following the QMRA framework, nor to quantification using empirical data. In the wastewater reuse domain, the scope of risk assessment could be extended to incorporate characterisation and optimisation of the treatment chain with prediction of effluent quality, or even, with an overarching sustainability focus, optimisation of effects on economic and environmental variables in addition to health-related outcomes.

5.5 CONCLUSION

The aim of this study was to illustrate the expediency of BNs for assessing and managing microbial risk. We have presented a BN based on a QMRA framework in a case study of wastewater reuse, parameterised using published data, which could be considered a prototype for future use of the method in water-related risk assessments. As indicated earlier, this BN for microbial health risk assessment could be extended and its utility enhanced by incorporation of other influences on health risk such as those posed by residual chemicals in treated wastewater. Due to the flexibility inherent in their design and quantification, BNs are an iterative tool that can be continually extended with more variables of diverse data types including expert opinion and/or updated with structural modifications or more exact data. They impart the benefit of additional uncertainty reduction resulting from each cycle of the knowledge engineering process. In a context of health risk assessment, BN's are therefore eminently suited to adaptive management and translational research.

Chapter 6: Potential of Bayesian networks for adaptive management in water recycling

Preamble

This chapter has been written as a journal article to meet Objective 4 of the research, as stated in the Introduction:

Objective 4 - To develop concurrent BNs representing the principal waterborne pathogen groups for water recycling and to validate their utility in assessment and management of wastewater treatment and reuse.

Following the development of the BN prototype described in Chapter 4, a set of three BN models representing the three key pathogen groups was conceived and planned for an alternative water reuse scenario, irrigation of public space. The model design incorporates four options for exposure profiles and the ability to choose two levels of a number of treatment chain steps and onsite reduction strategies, in combination.

This chapter is primarily my own work and the figures and tables were created by me. The article has been submitted to Environmental Modelling and Software and is reproduced here in its entirety.

Contributor	Statement of contribution
D. Beaudequin	Conception of models, conduct of research, model development and population, data analysis, interpretation of results, writing of manuscript, modifications to manuscript as suggested by co-authors and reviewers
F. Harden	Comments on manuscript, editing
A. Roiko	Comments on manuscript, editing
K. Mengersen	Comments on manuscript, editing

Principal Supervisor Confirmation: I have sighted email or other correspondence for all co-authors confirming their authorship.

Name: KERRIE MENGERSEN Signature: _____ Date: 14/3/16

Abstract

Water recycling is a valuable solution to increasing water demands and scarcity. However lack of data impedes fit-for-purpose water recycling and uniform wastewater treatment standards and health risk benchmarks deter uptake of water recycling schemes. Water managers and regulatory authorities are challenged by decision making under complex, uncertain conditions. Bayesian networks (BNs), a probabilistic risk assessment approach, are increasingly recognised as a valuable tool for adaptive management and decision making under uncertainty.

In this paper, we describe development and evaluation of a suite of three BNs for modelling health risk associated with wastewater irrigation of public open space. Concurrent BNs based on stochastic quantitative microbial risk assessment (QMRA) methods, representing the three major waterborne pathogen groups, are used to model multiple scenarios and exposure profiles. The BNs are designed to model risk reduction potential along a wastewater treatment chain as well as at the site of reuse and have the capacity to model a number of exposure profiles within a reuse scenario. The BNs provide an estimate of the conditional probability of infection or illness that can be compared directly with established health-based targets.

Study findings highlight the significant impact of post treatment risk mitigation on health risk outcomes, despite challenging conditions in the treatment chain. In the assessment and management of health risk related to water reuse, BNs provide a transparent, defensible evidence base for water resource managers and regulators to describe and quantify risk pathways, compare decision options and predict outcomes of management policies.

6.1 INTRODUCTION

Water is an increasingly valuable commodity worldwide. An exponentially rising global population with attendant food and water requirements, in addition to climate-related decrease of fresh water supplies in some areas, are contributing to growing water scarcity. Irrigation, which currently consumes 80% of the world's fresh water (UN-Water, 2014), is an important focus area for solutions.

Reuse of previously unutilised water resources such as treated wastewater is now being actively pursued to augment fresh water supplies. Despite this, concern regarding residual microbial contamination and difficulties with reliable and accurate determination of risk continue to inhibit its acceptance and implementation. Assessment of reclaimed water as being fit for its intended purpose should not be contingent solely on pathogen reduction benchmarks, as reuse-related health risk depends on a host of factors in the treatment-to-reuse chain (Keraita et al., 2010, World Health Organization, 2008). These include multiple variables in purpose-driven exposure pathways and post-treatment risk abatement measures, such as subsurface and drip irrigation systems, irrigation withholding periods and restriction of public access during and after irrigation events (NRMMC-EPHC-AHMC, 2006).

'Fit-for-purpose' wastewater reuse requires development of tools for rapid scenario assessment. Ideally, such tools should incorporate all of the key influences on health risk, including effects on treatment performance, fate and transport of pathogens during storage and distribution to the point of release, exposure pathways and impacts on dose-response relationships, including assessment of individual susceptibility. In the same way as a quantitative characterisation of treatment performance is essential, a quantitative understanding of exposure and dose-response scenarios is imperative. The ability to integrate the effects of these elements on health risk outcomes plausibly and with transparency is also necessary.

Wicked problems presented by complex environmental systems require holistic solutions. Systems thinking is increasingly being used to make reliable inferences to underpin decision making in integrated environmental modelling (Whelan et al., 2014). Sentinel authorities such as the United States Environmental Protection Agency have called for a paradigm shift from viewing health risk challenges 'water contaminant by water contaminant', to systems-based approaches (Anastas et al., 2010, Cohen Hubal et al., 2011). The authors of a paper on a screening-level

assessment of microbial risks from wastewater reuse note that the most important limitation in carrying out exposure analysis in quantitative microbial risk assessment (QMRA) is a lack of quantitative data on pathogens in water and their relative reduction at each stage of the treatment train (Petterson et al., 2001). Bayesian networks (BNs) offer not only a systems approach but also a number of other useful features to the characterisation of complex environmental risk assessments (Jensen and Nielsen, 2007, Pearl, 2000). BNs are also able to be quantified using diverse data types, including expert opinion, when insufficient or no empirical data exist.

This paper demonstrates use of BNs as an innovative technique to both visualise and quantify microbial exposure pathways, extending the means currently available to assess microbial health risks in water recycling schemes. With the overarching objective of providing an approach that more credibly represents microbial risks and to facilitate greater accuracy and science-based decision making with regards to fit-for-purpose wastewater treatment and reuse, this study had the multiple objectives of bridging the gap between microbial treatment performance measures and health-based targets, while incorporating the multiple barrier risk reduction paradigm in the determination of conditional probabilities for illness and infection. Building on our previous work, described in Beaudequin et al. (2016) and elsewhere, we present BNs for three reference pathogens in a context of recycled water irrigation of public open space.

We begin with background information on QMRA and BNs and describe the pathogens and exposure scenarios considered. In the Methods section we describe the case study in more detail, including the assumptions and data sources and the multiple infection barriers modelled. The QMRA modelling phase is explained, and we describe the construction of the BNs. In the Results section we present firstly, a comparison of the risks for four 'visitor' profiles by pathogen class. We then demonstrate the flexibility and expedience of BNs by modelling three multifaceted, theoretical scenarios, and conclude with results of a sensitivity analysis.

6.2 BACKGROUND

6.2.1 Quantitative microbial risk assessment

QMRA is a structured approach to the quantitative assessment of the likelihood and severity of potential adverse health outcomes associated with microbial

exposures. Based on a formal risk assessment framework (NRC, 1983, NRC, 2009) the integration of data with mathematical models is conducted in a four step process comprising hazard identification, exposure assessment, dose response assessment and risk characterisation (CAMRA, 2013a, USEPA-USDA/FSIS, 2012). In countries such as The Netherlands and Canada, QMRA has replaced the use of indicator-based approaches to regulation of drinking water quality (Smeets, 2013) and this is also the case for recycled water guidelines in Australia (Bichai and Smeets, 2013).

6.2.2 The multiple barrier approach to risk reduction

The multiple barrier approach to risk reduction, as recommended by the World Health Organisation (WHO) guidelines for wastewater use in agriculture (WHO, 2006), is widely recognised as a safe and efficient approach to managing risk posed by microbial contaminants in water supplies (Deere et al., 2001, Keraita et al., 2010, NHMRC, 2004) Under the multiple barrier paradigm, pathogen reduction in wastewater treatment is evaluated in combination with log reductions achieved through a range of other post-treatment health protection measures. By complementing pathogen removal during treatment, post-treatment health protection measures facilitate recycling schemes where lower levels of treatment can be used to reduce pathogen concentrations, in conjunction with higher levels of post-treatment risk reduction strategies. Such measures include allowing pathogen die-off between irrigation and contact, wastewater application techniques, human exposure control and crop restrictions and food processing in agricultural applications.

6.2.3 Bayesian networks

BNs are a powerful risk assessment tool, based on probability distributions, for reasoning under uncertainty (Jensen and Nielsen, 2007, Pearl, 1988). They are comprised of a set of random variables linked by arrows indicating causal relationships, thereby forming a directed acyclic graph. Variables or nodes are categorised by a set of user-defined ‘states’ that can be qualitative, discrete or continuous (e.g., ‘true/false’, ‘high/low’, ‘>5/≤ 5’). Each state is assigned a conditional probability, reflecting the strength of the causal links between the variables (Jensen and Nielsen, 2007). Conditional probabilities are derived from empirical data, models, simulations or the opinion of experts. Figure 6.1 is an example of a simple BN of two factors influencing *Cryptosporidium* oocyst concentration in primary treated wastewater. Nodes with no influences indicated by

incoming arrows are referred to as ‘root’ nodes; nodes with incoming arrows indicating influences of other variables, are referred to as ‘child’ nodes.

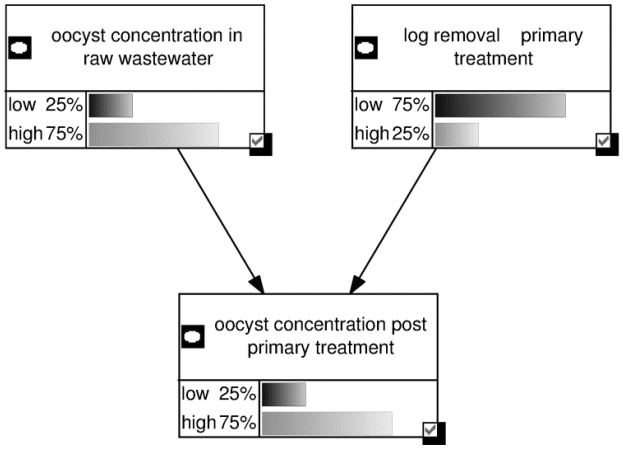

Figure 6.1. Simple Bayesian network showing causal influences on *Cryptosporidium* oocyst concentration in primary treated wastewater.

The conditional probabilities underlying child nodes in a BN are derived using Bayes' theorem, which evaluates the probability of an event, based on conditions that are thought to influence the event. Mathematically, Bayes’ theorem states:

$$P(B|A) = P(A|B)\ P(B)/P(A) \qquad (1),$$

where P(B|A) is the conditional probability of event B, or the probability of observing event B given that event A is true.

From probability theory it is known that P(A|B)P(B) = P (A and B). Thus the numerator in equation (1) is referred to as the joint probability of A and B, while the denominator is known as the marginal probability. The probability of an event B conditional upon another event A is calculated as joint probability/marginal probability.

An advantageous feature of BNs is that changes due to addition of new information or ‘evidence’ are immediately propagated through the network (Greiner et al., 2013). BNs are particularly useful for modelling and supporting decision making in complex systems such as are found in environmental domains. Although underpinned by complex algorithms, BNs are displayed via a visual platform and founded on basic probabilistic theory and are therefore highly accessible to non-statistical users and audiences.

6.2.4 Pathogens and exposure scenarios

QMRAs are typically limited to quantifying risk arising from exposure to a single pathogen. This is due to the pathogen-specific nature of the established dose-response relationships. Thus a pathogen-specific risk assessment in a water quality context does not reflect the reality that multiple pathogens may simultaneously present a threat to an exposed individual or population. This constraint can be addressed partially by undertaking a QMRA for a reference pathogen from each of the three significant pathogen groups (Mara, 2011). Bacteria, viruses and protozoa are represented in this study by *Campylobacter,* norovirus and *Cryptosporidium*, respectively. We examine these three organisms in a scenario of wastewater irrigation of a municipal park, sports field or golf course (hereinafter referred to as a park), in two potentially exposed populations: municipal employees and recreational users.

6.3 METHODS

6.3.1 Case study description

In this case study it is assumed that treated wastewater alone is used for irrigation of established grassed areas, delivered daily via spray nozzles. In the interests of a conservative assessment, it is assumed that no natural die-off or decline in infectivity occurs in pathogens as a result of desiccation or exposure to sunlight. Consideration of pathogen inactivation is limited to study of the effect of withholding irrigation prior to public access to the park. The exposure route studied is oral ingestion, occurring incidentally during work at, or a visit to the park, through direct hand-to-mouth contact or via fomites such as a golf ball or football.

It is assumed that golfers and casual park visitors are exposed to 1 mL of reclaimed water during a visit (Asano et al., 1992), and a person playing football is exposed to 5 mL of reclaimed water during a visit through ball handling and direct grass contact (Ryu, 2003). Occupational exposure is assumed to be 5 mL per day (Cherrie et al., 2006, Gorman et al., 2014), directly or indirectly, via contact with tools, equipment and green waste or inadvertent hand-to-mouth contact. Four ‘visitor’ profiles, summarised in Table 6.1, were modelled in the BNs by varying the volumes of wastewater ingested and the frequencies of wastewater contact, contingent on the activity undertaken. For example, infection risk for a golfer can be

simulated by setting the *Wastewater volume ingested* node to 1 mL (Asano et al., 1992, NRMMC-EPHC-AHMC, 2006) and the *Frequency of visits* to weekly. Similarly, exposure for a casual visitor to the park can be simulated by setting the *Wastewater volume ingested* node to 1 mL and the *Frequency of visits* node to fortnightly. Contact by a person playing football can be modelled by setting the *Wastewater volume ingested* node to 5 mL (Ryu, 2003) and the *Frequency of visits* to twice weekly and occupational exposure can be modelled by setting the *Wastewater volume ingested* node to 5 mL (Cherrie et al., 2006, Gorman et al., 2014), and the *Frequency of visits* to daily. Table 6.1 summarises the parameters of these visitor profiles and Tables 6.5, 6.6 and 6.7 in the Results section provides the baseline health risk estimates for the four profiles.

Table 6.1

Summary of simulated visitor profiles

visitor profile	wastewater volume ingested per visit (mL)	annual frequency of visits	reference
casual park visitor	1	fortnightly	(NRMMC-EPHC-AHMC, 2006)
golfer	1	weekly	(Asano et al., 1992, NRMMC-EPHC-AHMC, 2006)
football player	5	twice weekly	(Ryu, 2003)
occupational exposure (e.g., municipal worker)	5	daily	(Cherrie et al., 2006, Gorman et al., 2014)

The barriers to infection modelled in this study include six stages in a wastewater treatment chain and two post-treatment barriers. Two post-treatment barriers that have been shown to be effective in reducing public health risk in irrigation with recycled water are withholding irrigation (Page et al., 2014) and spray drift control (NRMMC-EPHC-AHMC, 2006). A recent study of microbial risk reduction by means of withholding periods during irrigation of public open space with recycled water (Page et al., 2014) has shown that a mean 0.7 $\log_{10}$ removal for bacteria and 0.4 $\log_{10}$ removal for viral pathogens can be achieved with a 4 hour withholding period after irrigation of sports ovals with secondary treated effluent. In

the present study *Cryptosporidium*, recognised as an 'environmentally recalcitrant' pathogen (Brooks et al., 2012), was assumed to undergo no appreciable inactivation in a 4 hour irrigation withholding period (Hutchison et al., 2005, Jenkins et al., 2013). Spray drift control, for which a 1 $\log_{10}$ reduction was assumed (NRMMC-EPHC-AHMC, 2006), can be achieved through measures such as low rise or low throw sprinklers, use of screening shrubs to intercept spray drift, irrigation at night and provision of personal protective equipment to workers exposed to recycled water (EPA QLD, 2005).

6.3.2 QMRA models

Deterministic QMRA models of risk of infection and illness for *Campylobacter jejuni*, norovirus and *Cryptosporidium parvum* were constructed initially in Microsoft Excel using point estimates from the literature. After validation, stochastic models were developed from the deterministic models, using triangular and uniform distributions with maxima and minima from the literature. QMRA model input parameters can be found in Tables E1, E3 and E5 in Appendix E. A dataset for each variable was generated by MC simulation with random sampling of 10,000 values from each distribution input using @Risk 6.2 (Palisade Corporation).

The pathogen concentration in treated wastewater at each stage of the treatment chain, C_{stage} (oocysts, PCR units or CFUs/mL for *Cryptosporidium,* norovirus or *Campylobacter* respectively), was modelled as

$$C_{stage} = 10^{(\log_{10} C_{previous} - LRV_{stage})} \quad (2)$$

(Water Research Australia, 2014b), where $C_{previous}$ is the pathogen concentration in wastewater from the previous step in the treatment chain and LRV_{stage} is the log removal of pathogens achieved through the method used in the current treatment stage ($\log_{10}$ units). The pathogen dose D (oocysts, PCR units or CFUs/mL) was modelled as

$$D = C_{onsite} V \quad (3),$$

where C_{onsite} is the final pathogen concentration at the point of exposure, after treatment chain pathogen reduction and onsite risk reduction strategies have been taken into account; V is the volume (mLs) of wastewater ingested by an individual in a visit to the wastewater-irrigated park.

As indicated previously, dose response is pathogen-specific and the models used are presented here:

Cryptosporidium

The risk of infection $P_{inf,}$ representing the risk of an individual contracting cryptosporidiosis as a result of a single visit to the park, was modelled as

$$P_{inf} = 1 - e^{(-kD)} \quad (4)$$

(Messner et al., 2001), where D is the pathogen dose and k = 5.72 x 10^{-2} (Center for Advancing Microbial Risk Assessment, 2013a).

Norovirus

The risk of infection $P_{inf,}$ representing the risk of an individual contracting a norovirus infection as a result of a single visit to the park, was modelled as

$$P_{inf} = 1 - \left(1 + \frac{D}{N_{50}}\left(2^{\frac{1}{\alpha}} - 1\right)\right)^{-\alpha} \quad (5)$$

(Teunis et al., 2008), where D is the pathogen dose, N_{50} is the median effective dose and α is the slope parameter. The values used for N_{50} and α were 1.7 x 10^4 and 1.11 x 10^{-1} respectively (Center for Advancing Microbial Risk Assessment, 2013a).

Campylobacter

The risk of infection $P_{inf,}$ representing the risk of an individual contracting campylobacteriosis as a result of a single visit to the park, was modelled as

$$P_{inf} = 1 - \left(1 + \frac{D_m}{N_{50}}\left(2^{\frac{1}{\alpha}} - 1\right)\right)^{-\alpha} \quad (6)$$

(Black et al., 1988), where D is the pathogen dose, N_{50} is the median effective dose and α is the slope parameter. The values used for N_{50} and α were 8.9 x 10^2 and 1.44 x 10^{-1} respectively (Center for Advancing Microbial Risk Assessment, 2013a). The remaining QMRA models are generic for the three pathogens:

The annual risk of infection $P_{inf(ann)}$was modelled as

$$P_{\mathrm{inf}(ann)} = 1 - \left(1 - P_{inf}\right)^{F} \quad (7)$$

(NRMMC-EPHC-AHMC, 2006), where F is the frequency of visits per year. The annual risk of illness $P_{ill(ann)}$ was modelled as

$$P_{ill(ann)} = P_{inf(ann)} R_{ill:inf} \quad (8)$$

(NRMMC-EPHC-AHMC, 2006), where $R_{ill:inf}$ denotes the illness to infection ratios for the respective pathogens.

The additional annual disease burden of illness, DB_{ann}, expressed in disability-adjusted life years per person per year (DALYs pppy), from each of the pathogens associated with visiting a wastewater-irrigated park, sports field or golf course was modelled as

$$DB_{ann} = DB_{case} P_{ill(ann)} S_f \quad (9)$$

(NRMMC-EPHC-AHMC, 2006), where DB_{case} is the disease burden per case (DALYs per case of pathogen-related illness) and S_f is the susceptibility fraction, or proportion of the population susceptible to the pathogen.

6.3.3 Bayesian network models

The structure of the BN models was based on the QMRA process models as described above, using GeNie 2.0 (University of Pittsburgh Decision Systems Laboratory, 2013). The structure of the BNs for the three reference pathogens was identical, except for *Cryptosporidium*, which did not have a *Withholding irrigation* node, due to assumptions regarding pathogen die-off discussed in section 6.3.2. Threshold values for node states were selected on the basis of equal probabilities in most cases, in the absence of relevant data (Tables E2, E4 and E6 in Appendix E). Other options for node discretisation could include using indicative or typical values for local conditions if such information was available, or equal bin size (Garcia et al., 2013).

The root nodes (nodes without parents) were parameterised using uniform probabilities where the prior distributions were unknown. Parameterisation of the states of child nodes was accomplished by calculating conditional probabilities from the simulated data set generated by 10,000 iterations of the stochastic process models. The BNs were evaluated through inference, scenario testing and sensitivity analysis.

The networks in Figures 6.2, 6.3 and 6.4 present the prior or 'baseline' state of knowledge of the risk of infection from each reference pathogen associated with visiting a park irrigated with reclaimed water, before any further specific evidence is

added. The value assigned to a node state represents the chance that the node will be in that state. For example, in the *Log removal primary treatment* nodes, in the absence of any other specific evidence, the chance of pathogen removal being low or high during primary treatment has been set at 50%.

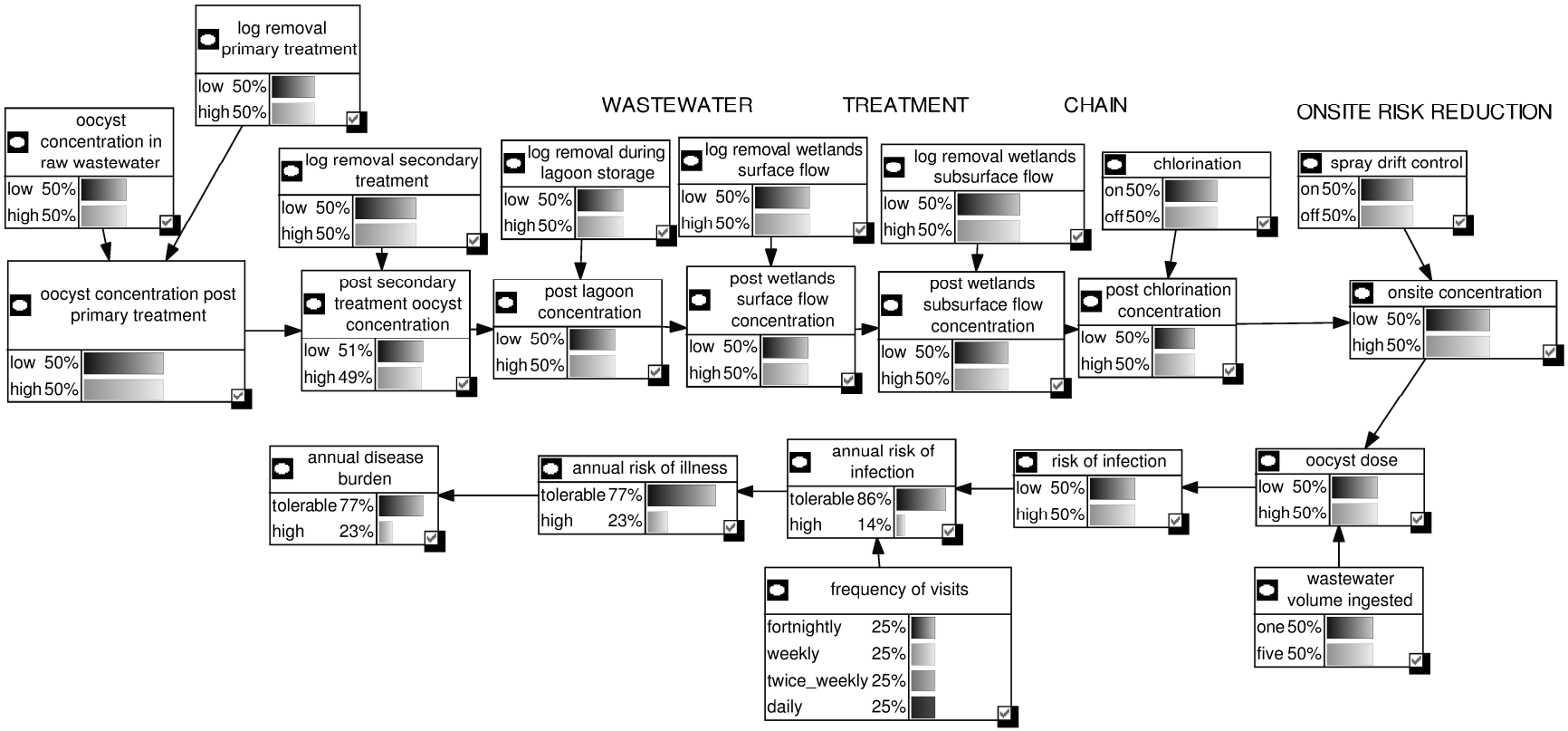

Figure 6.2. Risk of cryptosporidiosis as a result of visiting a park irrigated with reclaimed water.

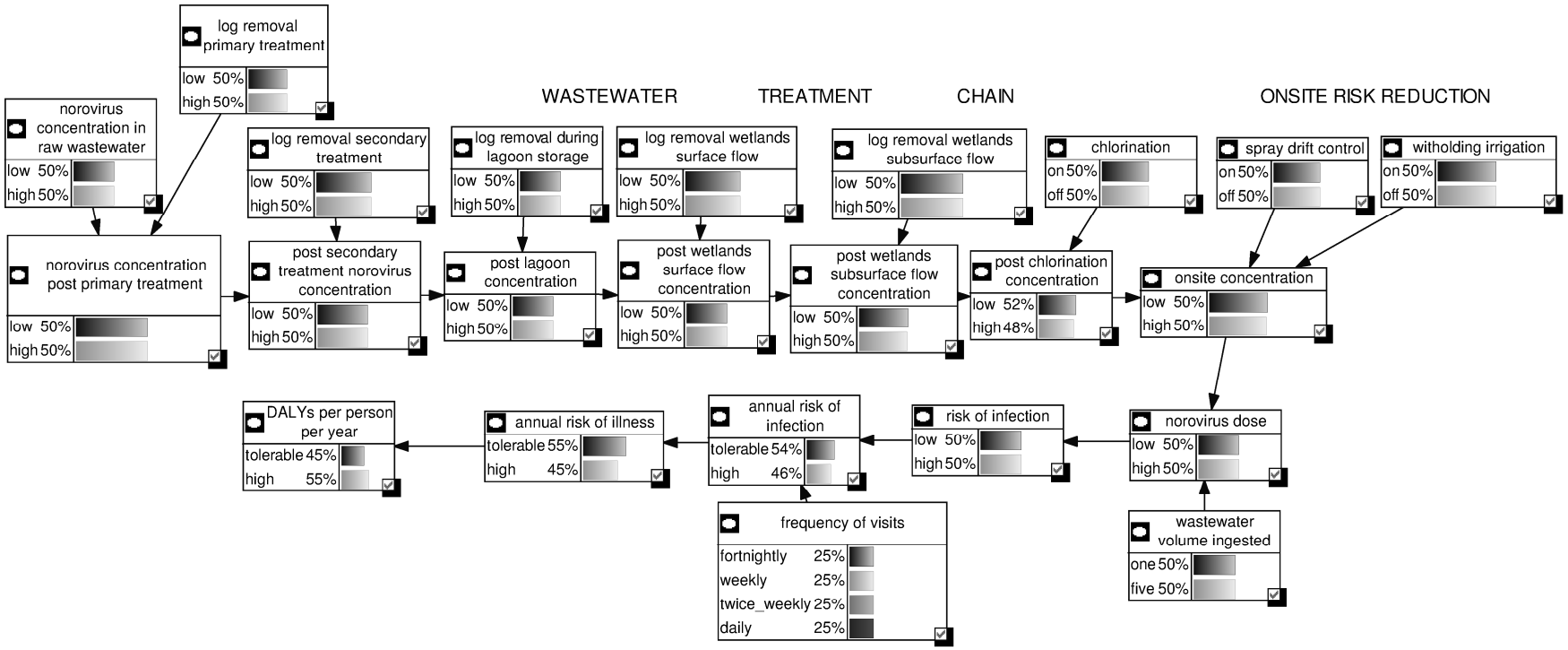

Figure 6.3. Risk of norovirus infection as a result of visiting a park irrigated with reclaimed water.

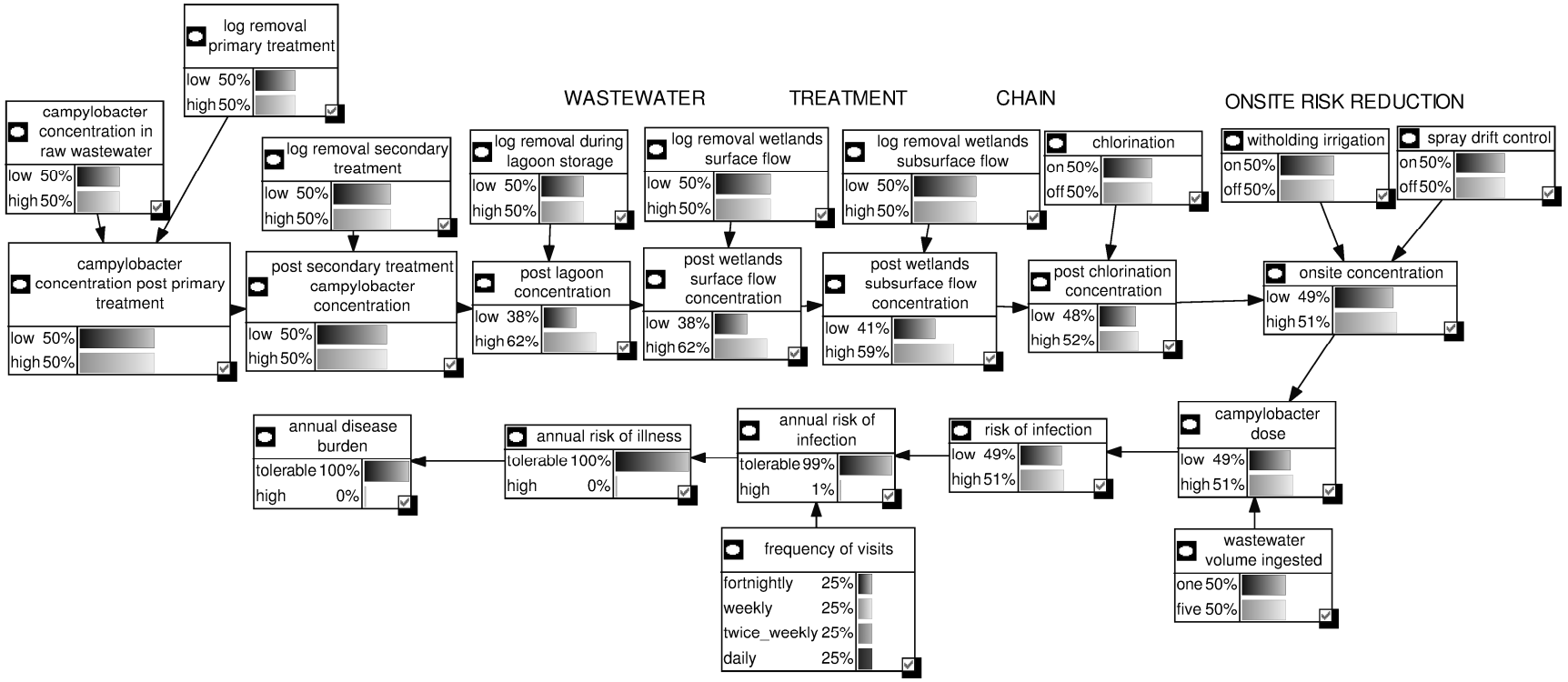

Figure 6.4: Risk of campylobacteriosis as a result of visiting a park irrigated with reclaimed water.

In QMRA terms, all nodes in the BNs (Figures 6.2-6.4) leading to and including the pathogen dose node contribute to characterisation of exposure. The *Risk of infection* node expresses the outcome of the pathogen-specific dose-response step and with the *Annual risk of infection*, *Annual risk of illness* and *Annual disease burden* nodes, represent the risk characterisation step of a QMRA. These are the four response nodes in the BNs. The *Annual risk of infection* node gives an annualised estimate of infection risk for an individual visiting a public park, golf course or sports field over a period of one year and correspondingly, the *Annual risk of illness* node provides information on the risk of illness for an individual visiting a public park, golf course or sports field over a period of one year. The *Annual disease burden* node estimates the chance of tolerable and high DALYs pppy, as a result of exposure to the treated wastewater in the scenarios modelled.

6.3.4 Derivation of conditional probabilities

Derivation of conditional probabilities for all possible state combinations in the nodes of the BNs began with discretisation of data into states using chosen threshold values, as shown in Table 6.2 for five iterations extracted from the full dataset:

Table 6.2

QMRA data for first three nodes of BN for cryptosporidiosis risk (Figures 6.1 and 6.2), discretised to states

iteration	oocyst concentration in raw wastewater (#/mL) low: ≤ 3.68 high: >3.68		log removal primary treatment low: ≤ 0.25 high: >0.25		oocyst concentration post primary treatment (#/mL) low: ≤ 2.04 high: >2.04	
1	3.15	L	0.29	H	1.59	L
2	6.41	H	0.01	L	6.31	H
3	4.28	H	0.27	H	2.28	H
4	3.22	L	0.29	H	1.64	L
5	5.09	H	0.44	H	1.87	L

As described earlier, equal probabilities were used for root nodes in the absence of data. Conditional probabilities for child nodes were calculated using equation (1) as follows: Table 6.3 is a contingency table for the child node *Oocyst concentration post primary treatment*. The table displays the number of times the states of *Oocyst concentration post primary treatment* occur with each of the states of its parent nodes *Log removal primary treatment* and *Oocyst concentration in raw wastewater*. For example, the number of instances of oocyst concentration in primary treated wastewater being high when log removal of oocysts during primary treatment was high and oocyst concentration in raw wastewater was also high, was 1697. The joint probability for this event is therefore 1697/10000 or 0.1697. The marginal probability is calculated as the number of instances when the *Oocyst concentration in raw wastewater* was high and the *Log removal primary treatment* was high (regardless of whether the *Oocyst concentration post primary treatment* was high or low): 2511/10000, or 0.2511. From equation (1), the conditional probability then, of *Oocyst concentration post primary treatment* being high when both influencing nodes are high, is calculated as the joint probability divided by the marginal probability, 0.1697/0.2511 = 0.6758. This probability is displayed in the bottom right hand cell in Table 6.4, which displays the full conditional probability table underlying the node *Oocyst concentration post primary treatment,* as an example of those underlying each node in the network.

Table 6.3

Contingency table for all possible state combinations for first three nodes of BN for cryptosporidiosis risk BN in Figures 6.1 and 6.2

Log removal primary treatment		**low**		**high**		
Oocyst concentration in raw wastewater		**low**	**high**	**low**	**high**	
Oocyst concentration post primary	low	1700	0	2489	814	
treatment	high	818	2482	0	1697	
grand total		2518	2482	2489	2511	10000

Table 6.4

Conditional probability table underlying the node Oocyst concentration post primary treatment node, from BN for cryptosporidiosis risk in Figures 6.1 and 6.2

Log removal primary treatment		**low**		**high**	
Oocyst concentration in raw wastewater		**low**	**high**	**low**	**high**
Oocyst concentration post primary	low	0.6751	0	1	0.3242
treatment	high	0.3249	1	0	0.6758

From Table 6.4 it can be seen that the probability of the oocyst concentration in primary treated wastewater being low when both the oocyst concentration in raw wastewater is low and the log removal of oocysts during primary treatment is also low, is 0.6751. Likewise, the probability of a high oocyst concentration in primary treated wastewater when the oocyst concentration in raw wastewater is high and log removal is also high, can be seen to be 0.3242. It can also be seen, reasonably, that it would be impossible ($P = 0$) to achieve a low oocyst concentration in the primary treated wastewater when the concentration in raw wastewater was high and the log removal was low and, as expected, a high concentration of oocysts in the primary treated wastewater is a certainty ($P = 1$) when their concentration in the raw wastewater is high and the log removal is low. The full BN for cryptosporidiosis risk associated with visiting a wastewater-irrigated park is shown in Figure 6.2. The BNs for norovirus infection and campylobacteriosis risk, as a result of visiting a park irrigated with reclaimed water, are displayed in Figure 6.3 and Figure 6.4, respectively.

6.3.5 Method of analysis

Transformation of node outcomes to a common metric for purposes of comparison can be achieved in different ways depending on the intention of the

model and influences of interest, or on the amount of variation from baseline in the target nodes. To study the effect of new evidence introduced into the network on the chosen health risk measures (such as setting nodes to certainty for specific states), the percent change in each response node state was calculated as

$$Change_{percent} = (P_{baseline} - P_{evidence})/P_{baseline} \times 100 \qquad (10),$$

where $P_{baseline}$ is the probability of occurrence of response node states under baseline network conditions before new evidence is introduced and $P_{evidence}$ is the probability of a state occurring after new evidence is introduced into the network.

The results of the BN modelling and analysis are presented in the following section for comparison of visitor profile risks, examination of risk under three hypothetical scenarios and a sensitivity analysis. The three hypothetical scenarios considered are:

1. Risk of norovirus infection on a golf course irrigated with recycled water under norovirus outbreak conditions;

2. High norovirus infection risk conditions for football players, with imposed constraint of 100% tolerable DALYs; and

3. High cryptosporidiosis risk conditions for municipal workers, with and without chlorination.

6.3.6 Sensitivity analysis

Sensitivity analysis of risk models can help identify the most significant factors to aid in risk management and is part of the model evaluation process. Analysis of sensitivity to evidence evaluates changes in the network in response to changes in inputs (Pollino and Henderson, 2010). Since the nodes in a BN do not exert equal influence on the response nodes, identification of the most influential factors on outcomes of interest can be used to prioritise further data collection for iterative model refinement in order to reduce uncertainty in risk estimates (Wang et al., 2002). The average sensitivity coefficients for the models in this study, presented in the following section, were computed by the method described by Kjaerulff and van der Gaag (2000).

6.4 RESULTS

6.4.1 Comparison of visitor profile risks

The BNs have been constructed with the ability to model four visitor profiles rapidly by manipulating the assumed frequency of visits and volume of wastewater ingested. Tables 6.5, 6.6, and 6.7 display baseline health risk measures for the visitor profiles for each reference pathogen, in the absence of any other evidence in the network, i.e., without introducing new evidence into nodes other than *Annual frequency of visits* and *Wastewater volume ingested.* As can be expected, the probability of achieving a tolerable risk decreases as the frequency of exposure and the volume ingested, represented by rank ordered visitor profiles, increase.

Table 6.5

Comparison of baseline response node probabilities for four visitor profiles – Cryptosporidium

visitor profile	annual risk of infection		annual risk of illness		DALYs per person per year	
	tolerable	high	tolerable	high	tolerable	high
casual park visitor	0.96	0.04	0.86	0.14	0.86	0.14
golfer	0.93	0.07	0.84	0.16	0.84	0.16
football player	0.80	0.20	0.72	0.28	0.72	0.28
municipal worker	0.66	0.34	0.60	0.40	0.60	0.40

Table 6.6

Comparison of baseline response node probabilities for four visitor profiles - norovirus

visitor profile	annual risk of infection		annual risk of illness		DALYs per person per year	
	tolerable	high	tolerable	high	tolerable	high
casual park visitor	0.70	0.30	0.71	0.29	0.58	0.42
golfer	0.65	0.35	0.66	0.34	0.54	0.46
football player	0.42	0.58	0.44	0.56	0.36	0.64
municipal worker	0.36	0.64	0.38	0.62	0.31	0.69

Table 6.7

Comparison of baseline response node probabilities for four visitor profiles – Campylobacter

visitor profile	annual risk of infection		annual risk of illness		DALYs per person per year	
	tolerable	high	tolerable	high	tolerable	high
casual park visitor	1.00	0	1.00	0	1.00	0
golfer	1.00	0	1.00	0	1.00	0
football player	0.99	0.01	1.00	0	1.00	0
municipal worker	0.98	0.02	0.99	0.01	0.99	0.01

In the following scenarios, the BN models were used to investigate the effect of risk reduction measures and other constraints under varying treatment chain conditions, on the health risk estimates in the response nodes.

6.4.2 Scenario 1: 'Norovirus outbreak'

A golf course is irrigated by treated wastewater from a treatment plant servicing a regional township. The plant is experiencing a high influent norovirus load due to an outbreak of norovirus in the town. The water utilities manager is not confident in the capacity of the primary and secondary treatment stages to remove viruses, due to its outmoded infrastructure, however the lagoon and wetlands system is modern and works well. The plant routinely uses chlorination, however facilities staff at the golf course do not generally use the onsite risk reduction measures spray drift control and withholding irrigation. To simulate these conditions, the following evidence is entered into the BN: *Norovirus concentration in raw wastewater* node is set to 100% high; *Log removal primary treatment* and *Log removal secondary treatment* nodes are set to 100% low; *Log removal during lagoon storage*, *Log removal wetlands surface flow* and *Log removal wetlands subsurface flow* nodes are set to 100% high; *Chlorination* node is set to 100% on; and the *Spray drift control* and *Withholding irrigation* nodes are set to 100% off. The BN in Figure 6.5 shows that the subsequent chance of a high risk of infection for an individual is 37%, while the chance of achieving tolerable DALYs is 57%.

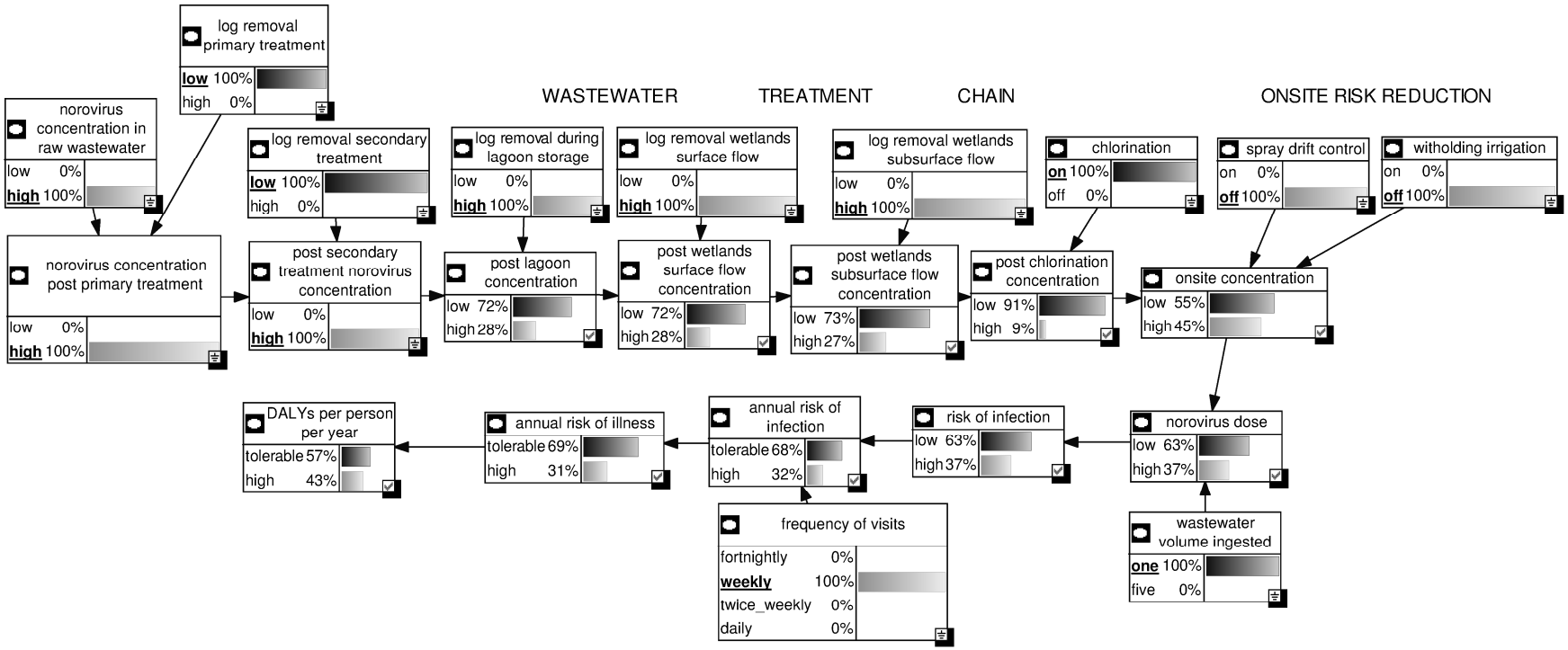

Figure 6.5. Scenario 1 - risk of norovirus infection on wastewater irrigated golf course under outbreak conditions with onsite risk reduction measures not in use.

Under the circumstances of the norovirus outbreak the water manager wants to demonstrate the efficacy of the onsite risk reduction strategies to golf course facilities staff. This is modelled by setting the *Spray drift control* and *Withholding irrigation* nodes to 100% on. The BN in Figure 6.6 illustrates the changed conditions. The chance of a high risk of infection has decreased to 5% and the updated chance of low DALYs - 78% - has increased significantly, by 37%. Response node changes for the scenario are summarised in Table 6.8. The most significant 'standardised' change in the response nodes was seen in the chance of a high annual risk of infection, which decreased by 88% when the onsite risk reduction measures were put into operation. The chance of a high annual disease burden was reduced by 49% as a result of using the onsite risk reduction measures spray drift control and withholding irrigation.

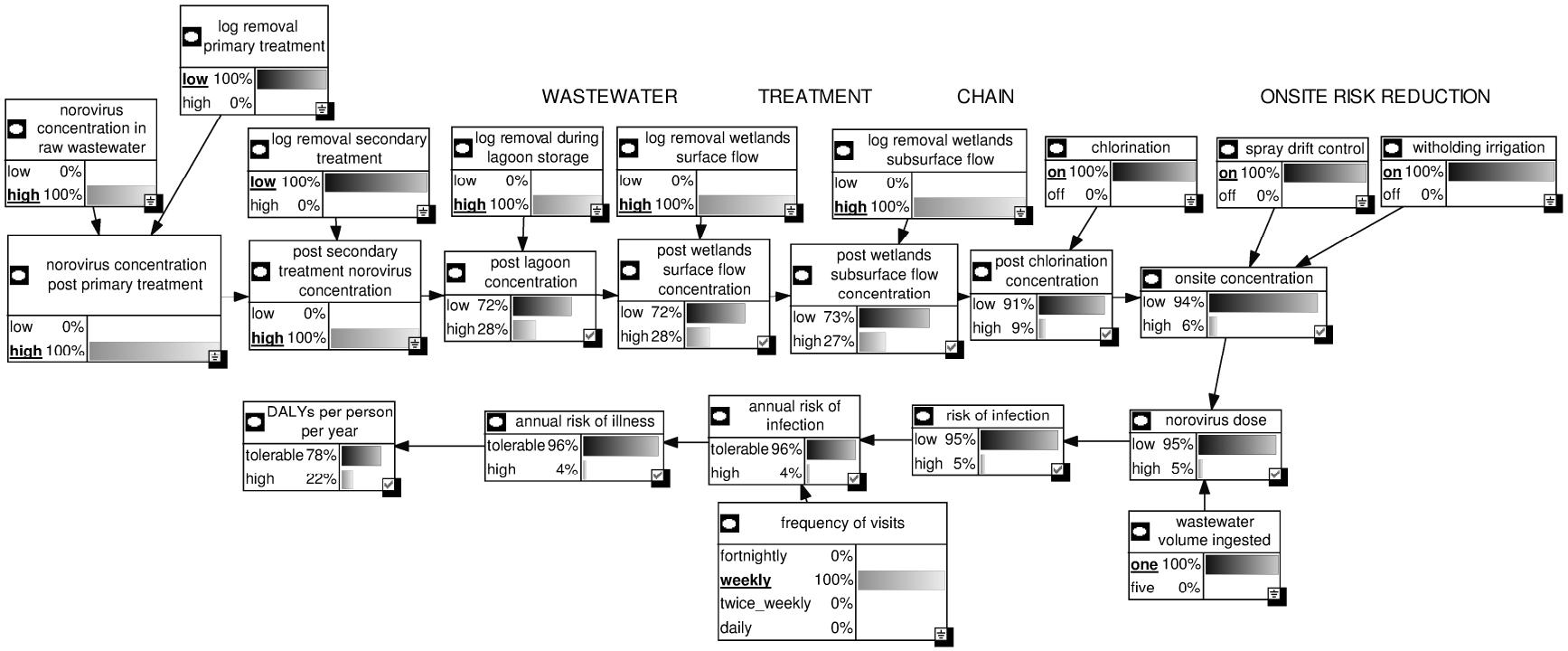

Figure 6.6. Scenario 1 - risk of norovirus infection under outbreak conditions with onsite risk reduction measures in use.

Table 6.8

Scenario 1 – norovirus infection risk for golf players. Chances of response node states with and without onsite risk reduction measures in operation

	annual risk of infection (%)		annual risk of illness (%)		annual disease burden (%)	
	tolerable	**high**	**tolerable**	**high**	**tolerable**	**high**
onsite risk reduction off	68	32	69	31	57	43
onsite risk reduction on	96	4	96	4	78	22
difference[a]	28	28	27	27	21	21
percent change[b]	41	88	39	87	37	49

[a]absolute value

[b]as discussed in the Method

6.4.3 Scenario 2: 'Certainty for tolerable burden of disease'

In this scenario a water utilities manager wants to assess the performance of the water treatment chain under maximum high infection risk conditions for suitability of irrigating a football oval, with the constraint that 100% tolerable DALYs is a certainty. These conditions are simulated in the network as shown in Figure F1 in Appendix F. This is an example of the ability of BNs to support 'backwards' inference, whereby the outcome required is entered in a response node and the network is examined to reveal the conditions of nodes supporting the required outcome. In this scenario, several challenging 'upstream' conditions are set in addition to the required outcome of 100% tolerable DALYs for norovirus infection: high concentration of pathogens in source waters, low log removals throughout the treatment chain and no chlorination or onsite risk reduction measures in use. If the conditions set are not achievable in the network, the software reports 'conflicting evidence'. In this case, despite challenging, high risk treatment chain conditions, the goal of 100% tolerable DALYs for footballers was achievable under high risk treatment chain conditions for norovirus. As expected, the high risk conditions in the

treatment chain resulted in a 99% chance of a football player receiving a high norovirus dose during a game, a 97% chance of high risk of infection and an 85% chance of high annual risk of infection. Nevertheless, the model indicated a comparatively smaller (21%) chance of high annual risk of illness, as norovirus infection does not necessarily result in illness, and also that tolerable disease burden (DALYS pppy) was achievable despite the conditions.

6.4.4 Scenario 3: 'Cryptosporidiosis risk'

A wastewater treatment plant in a rural catchment typically has a high level of *Cryptosporidium* oocysts in its source waters. The plant has very efficient primary and secondary treatment systems but a lagoon and wetlands system (surface and subsurface flows) which typically do not perform well as they are overrun by bird life. The treated wastewater is used to irrigate the municipal parks in an adjacent town, with spray drift control routinely used. The water utilities manager is under pressure to reduce costs and would like to know how important chlorination is for the control of cryptosporidiosis risk, given the high levels in source waters from surrounding cattle properties and the poor performance of the lagoon and wetlands system. The manager decides to evaluate the risk to municipal workers working daily in the town parks as a worst case exposure scenario, with and without chlorination. The BNs in Figures F2 and F3 in Appendix F shows network conditions for this scenario without chlorination and with chlorination, respectively. Table F1 in Appendix F shows the response node outcomes with and without chlorination, transformed to percent changes. The maximum effect of chlorinating the treated wastewater was a reduction of 6% in the chance of a high annual risk of infection for municipal workers and a reduction of 5% in the chance of a high annual disease burden, which the manager considered did not justify the expense of chlorinating before irrigation.

As an alternative to chlorination, the manager investigated the difference in health risk if the workers' daily visits were reduced to twice weekly. The BN for this change is shown in Figure F4 in Appendix F. Response node changes for the intervention are summarised in Table F2 in Appendix F. The effect was much more noteworthy than that of chlorinating the wastewater, with a 37% reduction in the chance of a high annual risk of infection and a 31% reduction in the chances of both a high annual risk of illness and high annual disease burden.

6.4.5 Sensitivity analysis

A useful property of BNs is the ability to identify dominant factors influencing nodes of interest, through sensitivity analysis. In the context of this study for example, sensitivity analysis is able to provide an indication of which treatment chain components and post treatment controls are best manipulated to produce maximum effect on a selected target node. Another application of sensitivity analysis is to provide information on which variables are critical areas for future research and data collection, to optimise research expenditure. In Tables 6.9 - 6.11, the relative strength of the association between the variable *Risk of infection* and its causal factors is indicated by the average sensitivity coefficient in the right hand column.

Table 6.9

Sensitivity analysis for risk of infection: Cryptosporidium

node	average sensitivity coefficient[a] for *Cryptosporidium: Risk of infection* node
Spray drift control	0.15
Log removal during lagoon storage	0.13
Chlorination	0.02
Log removal wetlands subsurface flow	0.02

[a] (Kjaerulff & van der Gaag, 2000; UP DSL, 2013)

Table 6.10

Sensitivity analysis for risk of infection: norovirus

node	average sensitivity coefficient[a] for norovirus: *Risk of infection* node
Chlorination	0.20
Spray drift control	0.11
Log removal during lagoon storage	0.08
Withholding irrigation	0.04

[a] (Kjaerulff & van der Gaag, 2000; UP DSL, 2013)

Table 6.11

Sensitivity analysis for risk of infection: Campylobacter

node	average sensitivity coefficient[a] for *Campylobacter: Risk of infection* node
Chlorination	0.33
Spray drift control	0.06
Withholding irrigation	0.04
Log removal during lagoon storage	0.04

[a] (Kjaerulff & van der Gaag, 2000; UP DSL, 2013)

Amalgamation of sensitivity coefficients for the three reference pathogens, including only those causal factors which can be manipulated by water treatment managers (i.e., disregarding factors such as wastewater volume ingested), resulted in the highest sensitivity factors for *Risk of infection* as shown in Table 6.12.

Table 6.12

Principal influences on Risk of infection node, ranked by sensitivity factor

node	average sensitivity coefficient[a]
Campylobacter: Chlorination	0.33
norovirus: *Chlorination*	0.20
Cryptosporidium: Spray drift control	0.15
Cryptosporidium: Log removal during lagoon storage	0.13
norovirus: *Spray drift control*	0.11
norovirus: *Log removal during lagoon storage*	0.08
Campylobacter: Spray drift control	0.06

[a] (Kjaerulff & van der Gaag, 2000; UP DSL, 2013)

6.5 DISCUSSION

6.5.1 QMRA results

Of the three reference pathogens modelled, norovirus had the lowest chance of achieving a tolerable annual disease burden, displaying the highest annual risk of infection and illness and highest DALYs estimate. Probabilities for *Annual disease burden* node in Figure 6.3 can be compared with corresponding values in Figure 6.2 and Figure 6.4. Norovirus is highly infectious (Ong, 2013), persistent in the environment (Ong, 2013; Silverman et al., 2013), resistant to wastewater treatment (Da Silva et al., 2008; NRMMC-EPHC-AHMC, 2006; Symonds et al., 2014) and is acknowledged as a challenging pathogen in the wastewater treatment domain. Synchronous risk assessments for viruses, bacteria and protozoa under the same exposure conditions in other investigations (Bastos et al., 2008, Mara et al., 2007, Pavione et al., 2013) also found viruses presented higher risks of infection than the other pathogen classes.

Annual disease burden estimates derived from the stochastic QMRA models (Table 6.13) show that health risk related to the pathogens spanned a range from 10^{-8} to 10^{-4} DALYs pppy, with mean values for *Cryptosporidium* and norovirus values exceeding the WHO (2006) guideline threshold of 10^{-6} DALYs pppy for acceptable level of risk from wastewater reuse (WHO, 2006). The median risk of infection for *Cryptosporidium* derived from the QMRA (1.84×10^{-6}) was of the same order of magnitude as that found in an analogous study of risk for human infection associated with wastewater irrigation of non-food crops (Carlander et al., 2009). The mean disease burden for *Campylobacter* (2.12×10^{-8} DALYs pppy) was two orders of magnitude lower than a corresponding mean disease burden for *Campylobacter* of

4.85 x 10^{-6} DALYs pppy for all withholding periods found in a recent study of irrigation of public open space irrigation with recycled water (Page et al., 2014). The same study, using rotavirus as the reference pathogen, reported a mean viral disease burden of 4.9 x 10^{-6} DALYs pppy (Page et al., 2014); the norovirus disease burden of 1.02 x 10^{-4} DALYs pppy in the present study was two orders of magnitude lower.

Table 6.13

Annual risks of infection and illness and DALYs for three pathogens from QMRA process models, with published respective tolerable values

pathogen	annual risk of infection	tolerable annual risk of infection	annual risk of illness	tolerable annual risk of illness	DALYs per person per year	tolerable DALYs
Cryptosporidium	2.03 x 10^{-3}	**2.2 x 10^{-3}	1.42 x 10^{-3}	**6.7 x 10^{-4}	2.14 x 10^{-6}	*1 x 10^{-6}
norovirus	5.02 x 10^{-2}	***1.4 x 10^{-3}	3.36 x 10^{-2}	***1.1 x 10^{-3}	1.02 x 10^{-4}	*1 x 10^{-6}
Campylobacter	1.46 x 10^{-5}	**3.2 x 10^{-4}	4.39 x 10^{-6}	**2.2 x 10^{-4}	2.12 x 10^{-8}	*1 x 10^{-6}

*(World Health Organisation, 2006b)

**(World Health Organisation, 2006a)

***(Mara & Sleigh, 2010)

6.5.2 BN for campylobacteriosis risk

The BN for *Campylobacter* indicated chances of 99%, 100% and 100% for achieving a tolerable annual risk of infection, tolerable annual risk of illness and tolerable DALYs, respectively (Figure 6.4). These results seemed incongruently high compared with the chances of achieving analogous targets generated for *Cryptosporidium* (Figure 6.2) and norovirus (Figure 6.3). *Campylobacter* concentration in raw sewage was modelled as a triangular distribution with a range of 0.1 – 100 CFU/mL as quoted in the Australian Guidelines for Water Recycling (NRMMC-EPHC-AHMC, 2006). However a report on the presence and removal of enteric pathogens in South East Queensland wastewater treatment plants described significantly higher *Campylobacter* densities of the order of 7 $\log_{10}$ CFU/100mL in raw sewage (Toze et al., 2012). This document described numbers at the genus level which included species other than *C. jejuni,* whereas the Australian guidelines reported numbers specifically for *C. jejuni*. It was also noted that although the numbers in the guidelines are taken from scientific literature, they are predominantly sourced from studies undertaken in North American wastewater treatment plants and

are therefore possibly lower than may be found in Australian conditions. It was also noted that the numbers for south east Queensland were higher than anticipated.

6.5.3 Health-based targets

The apparently discordant findings in Scenario 2 - a low (15%) chance of tolerable annual risk of norovirus infection and a 79% chance of tolerable annual risk of illness, in conjunction with 100% tolerable DALYs for football players – is an acknowledged limitation of current health benchmarks. The potential for achieving conflicting conclusions when two or more QMRA model endpoints (i.e., risk of infection, annual risk of infection) are reported together for the same scenario has been noted elsewhere (Barker, 2014b). There are a number of points to be made in this regard. Commonly used tolerable or acceptable risk benchmarks have been described as 'relative and judgemental' quantities (Sinclair et al., 2015), albeit usually based on best available scientific evidence and opinion. However their inherent assumptions may not be a direct match for the conditions being modelled; in these respects they can be viewed as being both subjective and variable. A second consideration in using tolerable risk benchmarks is that they are usually cited as point estimates, without indication of the degree of uncertainty attending them. Thirdly and most significantly, when tolerable risk benchmarks are used as threshold values in BNs, model conclusions are not translated dichotomously as 'safe' or 'unsafe', but are presented as a continuous scale of chance of achieving target levels, as seen in this study. Transformation of risk benchmarks to a probability continuum is another positive feature of the BN methodology.

6.5.4 Static and dynamic risk assessments

The models in this study represent a static risk assessment, assessing individual risk as opposed to population-based risk, considering only direct or environment-to-person exposure and disregarding the potential for immunity or other influences on susceptibility to infection in individuals (Beaudequin et al., 2015b). More sophisticated, dynamic models are not necessarily more accurate, as risk levels of concern decrease, the estimates from static and dynamic models have been shown to converge (Soller et al., 2004). Nevertheless, in some cases consideration of the potential for secondary transmission and immunity in a dynamic risk assessment may produce a more accurate representation of risk. In dynamic risk assessment for instance, all exposed individuals may not be susceptible to infection as they may

already be infected or immune from prior exposure. This status changes with time as the infected overcome the pathogen and as immunity wanes. By taking into account indirect (person-to-person) exposure routes and the possibility for immunity, dynamic or time-varying risk assessments consider the movement of individuals in and out of susceptible, infected and immune states over time. The incorporation of time as a variable can be achieved with object-oriented and dynamic BNs (Johnson et al., 2010, Johnson and Mengersen, 2011). BN-based risk assessments founded on individual exposure scenarios can also be extrapolated to population-based assessments through incorporation of estimates of the resident local population, age groups or facility users.

6.5.5 Potential variations to design

The potential for the BN methodology in this domain is significantly broader than can be conveyed here. The design of these networks could be varied in several ways to reflect an integrated environmental modelling framework (Whelan et al., 2014). Node states are not limited to Boolean categories but can be adapted to the problem being modelled; for instance, the pathogen concentration in wastewater could be modelled at higher resolution using more states; or a BN might include other variables such as water class A to D (EPA, 2005), nominal age groups for exposed individuals, or time of day categories. Pathogen concentration is arguably an important influence on health risk and influences on this variable in a wastewater context might include variations in the treatment chain not represented in this study For example, experiential data on operating conditions, influent loads (Li et al., 2010, Li et al., 2013) and environmental factors (Beaudequin et al., 2015b), as well as influences during post-treatment transport, storage and distribution. Other treatment-related parameters, such as disinfectant, lagoon detention time, biochemical oxygen demand or suspended solids could be included as surrogates for treatment performance relating to pathogen reduction (Beaudequin et al., 2015b, NRMMC-EPHC-AHMC, 2006).

Microbial indicator data could also be used as BN inputs in this framework. While advisability of the sole use of bacterial indicator organisms such as *E.coli* in the microbial assessment of water quality is questioned by many authors (Alcalde et al., 2012, Harwood et al., 2005, O'Toole et al., 2014, Payment and Locas, 2011, Wu et al., 2011), waterborne pathogens continue to be both difficult and costly to

enumerate and pathogen monitoring for determination of water quality is often beyond the means of water authorities in developing countries. In the absence of other data, microbial water quality indicators such as *E. coli* or faecal coliforms can be incorporated into BNs through use of established ratios as has been done in some QMRAs (Mara and Sleigh, 2010, Mara et al., 2007, Seidu et al., 2008). In this cases assumptions and conditions underlying development of ratios, such as study location, need to be closely examined for fit to the conditions being modelled (O'Toole et al., 2014). Correlation and prediction relationships such as regression have also been used to estimate health risk of pathogens from indicators, as was done in a BN based on QMRA for cryptosporidiosis and giardiasis in small private water supplies (Hunter et al., 2011). In addition to components of these components of the exposure pathway, other elements with strong influences on dose response (Beaudequin et al., 2015b) may also be incorporated with relative ease in a BN despite a lack of data.

6.5.6 Potential variations to method

In addition to variations to BN structure, the way in which conditional probabilities are elicited could also vary. The conditional probabilities for the BNs in this study were derived from data sets generated by MC simulation in stochastic QMRA models, with inputs for the QMRA models obtained from the literature. However BN inputs could be obtained from field work or conditional probabilities could be estimated directly by expert teams (O'Hagan et al., 2006), without recourse to QMRA modelling. BNs can also be quantified directly from datasets using inbuilt algorithms (Fenton and Neil, 2013). Alternatively, nodes in a BN can be populated in different ways using multiple methods. Threshold values for node states are another feature of BNs which can be changed to meet established targets, or to reflect inclination for risk; for instance, avoiding the highest risk or achieving the lowest risk. If no empirical data are available to guide the application of relevant thresholds, equal width or equal frequency are valid discretisation methods yielding a useful model (Garcia et al., 2013). Threshold values are an explicit component of BNs, as once set and modelled they are generally displayed in the supporting documentation. These potential variations to the BNs developed for this study serve to illustrate the convenient plasticity of the BN methodology.

6.5.7 Concurrent microbial exposures

As previously indicated, QMRA is necessarily a pathogen-specific risk assessment method. Presentation of concomitant QMRA models in BN format for the three representative pathogen classes of significance in waterborne disease provides a convenient synchronistic overview of the relative effects of in-treatment and post-treatment influences on health risk associated with each pathogen class. The synchronised consideration of the three reference pathogens under the same exposure conditions, with the clear result of norovirus being the most likely to cause infection, adds weight to the mounting case for viral indicators of microbial water quality (Kitajima et al., 2014, Mok and Hamilton, 2014). Contemporaneous consideration of multiple pathogens in the same exposure pathway also raises the question of the effect of concurrent pathogenic exposures on the individual, since contact with wastewater potentially entails simultaneous introduction of a range of microorganisms, pathogenic and otherwise, to the host. In addition to colonisation, invasion, multiplication and antibody response, interspecific population dynamics such as competition, predation and mutualism or synergism could theoretically take place within the host. The outcome may be infection of one tissue type or organ by one pathogen type, or by multiple pathogen types, or infection in multiple tissue or organ types by multiple types of pathogens. Exploration of the immune response in concurrent pathogenic exposure is beyond the scope of this paper, however passing reference is made here due to its relevance to the treatment of the multiple health risk estimates generated by concomitant pathogen models. The scarcity of literature on the topic indicates that much needs to be accomplished in order to understand (and predict) concurrent pathogenic exposures.

6.6 CONCLUSION

The aim of this study was to illustrate the utility of using BNs to evaluate and analyse influences in the exposure pathway of a microbial health risk assessment. The models developed and discussed in this paper demonstrate the unique way in which BNs can be used to aggregate QMRA with other types of information and in which the BN modelling approach characterises causal relationships, identifies key influences on outcomes of interest and reveals significant knowledge gaps in the exposure pathway. In the combination of the BN and QMRA approaches, the study demonstrates the successful incorporation of explicit uncertainty into QMRA,

extending research methods and expanding knowledge in the interests of public health. In conjunction with other tools such as disease surveillance and epidemiological studies in the risk assessment arsenal, the BN methodology has considerable potential to overcome data limitations and other constraints in the study of environmental exposures to microbiological organisms.

Chapter 7: Discussion

Environmental exposures are typically difficult to characterise, due to the vagaries of ecological influences on the contaminant and the complexity of the contaminant-receptor and receptor-environment interactions. Microorganisms as contaminants add another layer of complexity to an already challenging domain, as issues such as enumeration, microbial population dynamics, person-to-person transmission, immunity and susceptibility further confound risk characterisation. BNs are a powerful, complementary method of addressing the complexities of microbial exposures in environmental domains. This thesis facilitates fit-for-purpose wastewater reuse, through development, evaluation and promulgation of BN modelling in assessment of risk associated with microbial pathogens.

The series of four publications that comprise the body of the research address the knowledge gaps outlined in the Introduction, accomplishing the four objectives of the research program. In Chapter 3 the literature on the previous use of BNs in microbial risk assessment at large was explored, to identify and examine what had been achieved in this niche area. Although BNs had not been extensively used in conjunction with QMRA at the time of writing, a major finding was widespread endorsement by authors of their benefits in microbial risk assessment, with due reference to drawbacks and limitations of the method. BNs were generally found to provide a number of attributes that simplified scenario analysis and overcame many of the acknowledged limitations in established QMRA methods. A drawback highlighted by some authors was the difficulty with quantifying conditional probability tables; however the review also revealed that one of the aspects of flexibility in BNs is the number of methods of eliciting probabilities to parameterise the nodes. The review confirmed that BN applications in the investigation of microbial health risk in the wastewater reuse domain is relatively novel. Perhaps a manifestation of an emerging field, the papers reviewed were notable in the disparity of the technical language used and transparency of the modelling process.

The development of this novel application of BNs progressed with the conceptualisation of health risk modelling in a water recycling context. Development

of a conceptual model established the foundation for the refined, parameterised networks to follow. Chapter 4 described this initial stage of the modelling process, development of an unparameterised causal network elicited by experts and substantiated by relevant literature. The conceptual model comprised four submodels, of which three (the *Exposure, Dose response* and *Risk characterisation* submodels) were generic to water or wastewater exposure scenarios and one (the *Pond operation and performance* submodel) explored potential influences on pathogen concentration in a specific wastewater treatment step. Risk reduction strategies, a function of risk management, were not included in this conceptual model of risk assessment, but were discussed as a foundation for inclusion in the subsequent quantified models integrating risk assessment and risk management options. This work resulted in an increase in the clarity and accessibility of QMRA steps.

Chapter 5 subsequently described development and evaluation of a refined, parameterised BN. The model, based on established QMRA modelling procedures, quantified the health risk of a common wastewater reuse scenario, consumption of wastewater-irrigated lettuce. The purpose of the study described in this chapter was to demonstrate the convenience of rapid assessment of multiple exposure and intervention scenarios, alone and in combination and to examine effects on health risk estimates. Table 7.1 provides a summary of the scenarios modelled with this BN and subsequent effects on the chance of a low risk of infection.

Table 7.1

Bayesian network for norovirus infection associated with wastewater irrigated lettuce: summary of scenario outcomes described in Chapter 5

scenario	evidence	chance of low risk of infection (% difference from baseline*)
baseline	none	59% (baseline)
'Outbreak'	*Norovirus concentration in treated wastewater*: 100% high	48% (-19%)
'Outbreak with risk mitigation'	*Norovirus concentration in treated wastewater*: 100% high; *Irrigation withholding period*: 100% high	82% (+39%)
'Furrow system'	*Wastewater retained on lettuce*: 100% low	62% (+5%)
'Treatment change'	*Norovirus concentration in treated wastewater:* 100% low	68% (+15%)
'Lettuce washing'	*Pathogen reduction (washing):* 100% high	77% (+31%)
'Rain'	*Norovirus concentration in treated wastewater*: 100% high; *Wastewater retained on lettuce*: 100% low	50% (-15%)
'Rain with decreased withholding period'	*Norovirus concentration in treated wastewater:* 100% high; *Wastewater retained on lettuce*: 100% low; *Irrigation withholding period:* 100% low	19% (-67%)

*calculated as (baseline value – changed value)/baseline value x 100

Sensitivity analysis showed that risk reduction strategies of withholding irrigation for 3 days prior to harvesting and lettuce washing had the most influence on risk of infection, followed by pathogen concentration in treated wastewater. A three day irrigation withholding period resulted in a 94% reduction in the chance of a high risk of infection, which was the equivalent of doubling the chance of meeting the targets for tolerable infection risk. Lettuce washing doubled the chance of achieving a tolerable annual risk of infection. Use of maximum post-treatment risk controls together resulted in a high probability (0.84) of low risk of infection. A notable finding was that with maximum risk controls in use, all three health risk outcomes were unaffected by the introduction of new evidence in other variables, including changing the *Norovirus concentration in treated wastewater* to high.

In Chapter 6 a second application of the BN methodology enabled the flexibility of BNs to be further explored and demonstrated, in the assessment and management of risk related to wastewater treatment and exposure scenarios. This phase of the research also presented the opportunity to investigate the value of developing concurrent networks representing three principal pathogen groups of concern in waterborne disease – bacteria, viruses and protozoa. Furthermore, building on the utility of the network described in the previous chapter, the BNs in Chapter 6 were designed to model risk reduction potential along a wastewater treatment chain as well as at the site of use. A third feature of the BNs in this chapter was the capacity to model a number of exposure profiles within a reuse scenario. Table 7.2 summarises the health risk outcomes for scenarios modelled in this study.

Table 7.2

Bayesian networks for norovirus infection and cryptosporidiosis risk, associated with wastewater irrigation with public open space: summary of scenario outcomes described in Chapter 6

scenario	**evidence**		**chance of tolerable annual burden of disease (% difference from baseline*)**
Norovirus outbreak risk to golf players, onsite risk reduction not in use	*Wastewater volume ingested*	100% one	57% (baseline)
	Frequency of visits	100% weekly	
	Norovirus concentration in raw wastewater	100% high	
	Log removal primary treatment	100% low	
	Log removal secondary treatment	100% low	
	Log removal during lagoon storage	100% high	
	Log removal wetlands surface flow	100% high	
	Log removal wetlands subsurface flow	100% high	
	Chlorination	100% on	
	Spray drift control	100% off	
	Withholding irrigation	100% off	
Norovirus outbreak risk to golf players, onsite risk reduction in use	*Wastewater volume ingested*	100% one	78% (+37%)
	Frequency of visits	100% weekly	
	Norovirus concentration in raw wastewater	100% high	
	Log removal primary treatment	100% low	
	Log removal secondary treatment	100% low	
	Log removal during lagoon storage	100% high	
	Log removal wetlands surface flow	100% high	
	Log removal wetlands subsurface flow	100% high	
	Chlorination	100% on	
	Spray drift control	100% on	
	Withholding irrigation	100% on	
Certainty of tolerable burden of disease for norovirus infection in football players	*Wastewater volume ingested*	100% five	100% (certainty)
	Frequency of visits	100% twice weekly	
	Norovirus concentration in raw wastewater	100% high	
	Log removal primary treatment	100% low	
	Log removal secondary treatment	100% low	
	Log removal during lagoon storage	100% low	
	Log removal wetlands surface flow	100% low	
	Log removal wetlands subsurface flow	100% low	
	Chlorination	100% off	
	Spray drift control	100% off	
	Withholding irrigation	100% off	
	Annual burden of disease	100% tolerable	
Cryptosporidiosis risk to municipal workers – chlorination not in use, daily exposure	*Wastewater volume ingested*	100% five	58% (baseline)
	Frequency of visits	100% daily	
	Oocyst concentration in raw wastewater	100% high	
	Log removal primary treatment	100% high	
	Log removal secondary treatment	100% high	
	Log removal during lagoon storage	100% low	
	Log removal wetlands surface flow	100% low	
	Log removal wetlands subsurface flow	100% low	
	Chlorination	100% off	
	Spray drift control	100% on	

scenario	evidence		chance of tolerable annual burden of disease (% difference from baseline*)
Cryptosporidiosis risk to municipal workers – chlorination in use, daily exposure	*Wastewater volume ingested*	100% five	60% (+3%)
	Frequency of visits	100% daily	
	Oocyst concentration in raw wastewater	100% high	
	Log removal primary treatment	100% high	
	Log removal secondary treatment	100% high	
	Log removal during lagoon storage	100% low	
	Log removal wetlands surface flow	100% low	
	Log removal wetlands subsurface flow	100% low	
	Chlorination	100% on	
	Spray drift control	100% on	
Cryptosporidiosis risk to municipal workers – chlorination not in use, twice weekly exposure	*Wastewater volume ingested*	100% five	71% (+22%)
	Frequency of visits	100% twice weekly	
	Oocyst concentration in raw wastewater	100% high	
	Log removal primary treatment	100% high	
	Log removal secondary treatment	100% high	
	Log removal during lagoon storage	100% low	
	Log removal wetlands surface flow	100% low	
	Log removal wetlands subsurface flow	100% low	
	Chlorination	100% off	
	Spray drift control	100% on	

*calculated as (baseline value – changed value)/baseline value x 100

A key finding in this study was that use of onsite risk reduction measures - spray drift control and withholding irrigation for four hours before public contact – achieved a 37% increase in the chance of meeting the tolerable annual burden of disease benchmark for norovirus despite treatment chain performance anomalies. The BN for cryptosporidiosis risk in municipal workers also demonstrated a notable positive effect on health risk outcomes with the simple measure of reducing exposure frequency, resulting in a 22% increase in chance of achieving a tolerable burden of disease for municipal workers. This result was in distinct contrast to the relatively slight effect of chlorination (3% increase in chance of tolerable burden of disease) on reducing risk when treatment chain performance was suboptimal. Study findings in both Chapter 5 and Chapter 6 highlight the significant impact of post treatment risk mitigation on health risk outcomes, despite challenging conditions in other variables in the exposure pathway, including pathogen concentration in raw or treated wastewater.

Development and demonstration of BN faculties in the studies in this thesis is a small indication of the power of BNs and strengthens the argument for their value in

assessing and managing complex systems. The BN modelling method used throughout this work unquestionably accommodates the National Research Council's (2009) contemporary risk assessment framework outlined in Chapter 2, wherein the question is asked:

"What options are there to reduce the hazards or exposures that have been identified and how can risk assessment be used to evaluate the various options?"

When numerous factors are thought to influence an outcome there is often a natural inclination to include as many factors as possible, making a model overly complex. Although larger models with increased complexity imply greater precision and may inspire more confidence than simpler models, accuracy may not necessarily be improved, due to amplification of uncertainty from additional parameters (Zwietering, 2009). To identify conditions under which static and dynamic models yielded significantly different results, Soller and Eisenberg (2008) carried out a study comparing a static, individual-level risk model to a dynamic, population-level model that included secondary transmission and immunity processes, based on a scenario of human pathogen exposure associated with reclaimed water. They concluded that under low risk conditions, the simpler static model provided satisfactory risk estimates (Soller and Eisenberg, 2008). Simple and complex approaches have their place in risk assessment, with simple models providing both insight and serving to detect major factors and potential errors in more complex models (Zwietering, 2009). It has also been argued that whilst errors may also abound in simple models due to oversimplification, simple modelling approaches carry benefits of increased transparency, practicality and availability of parameters and 'the domain of validity of the simpler approach can be investigated using the complex approach' (Zwietering, 2009).

As discussed in Chapter 2, adaptive management is based on an iterative decision making, monitoring and learning cycle, improving long term management outcomes through making short term decisions, observing the outcomes and modifying management strategies as understanding of the system improves (Holling, 1978, Walters, 1986). Similar to the 'plan-do-check-act' quality improvement method used in business for control and continuous improvement of processes and products (Walton and Deming, 1986), adaptive management brings about robust decision making in the face of commonly encountered uncertainty in environmental

domains. Instead of using a single set of probability distributions, adaptive management strategies use multiple representations of the future, or scenarios, to characterise and reduce uncertainty (Lempert and Collins, 2007). BNs are reported to be beneficial in adaptive management approaches, as they support rapid 'what if' analyses and iterative improvement methods (Barton et al., 2012, Landuyt et al., 2013, Pollino and Henderson, 2010). The studies in Chapter 5 and 6 clearly validate the use of BNs for adaptive management in water recycling, through the iterative nature of the knowledge engineering process integral to BN development and through their efficient support of scenario analysis in the characterisation and reduction of uncertainty (Lempert and Collins, 2007). This use of the BN methodology also supports the call for a more pre-emptive approach in the identification and management of risk in water reuse programs, as opposed to depending solely on posttreatment testing for managing risk (Mok and Hamilton, 2014).

These studies clearly demonstrate the appealing features of BN models in a water recycling context. The respects in which the technique complements QMRA, rendering it more transparent, flexible and accessible to stakeholders and decision makers are summarised here.

- **BNs are able to support 'backwards' inference,** enabling discovery of the key drivers for an outcome of interest. Alternatively, the desired outcome can be introduced as evidence in a target node, to determine conditions required 'upstream' to achieve the desired outcome.
- **Complex scenario analysis is simplified using BNs**, as new evidence can be introduced to multiple variables to simulate changed system conditions.
- **BNs have instant updating capability**, providing rapid results for scenario appraisal. The influence of variable changes on the joint distribution is propagated through the network instantaneously by software algorithms and resulting changes in outcome nodes are immediately visible.
- **The influence of new evidence is propagated throughout the network in both directions,** which also results in updating of poor quality prior information. This is particularly useful in QMRA, where data may be poor in quality due to the inherent difficulties of microorganism enumeration.

- **BNs represent uncertainty clearly** at variable level, through graphic probability distributions.
- **BNs transform current risk benchmarks** from a dichotomous structure (e.g.,'tolerable/not tolerable') to a probability continuum (e.g., '94% chance of achieving tolerable risk of infection').
- **BNs are visual models,** promoting engagement and enabling stakeholders from different disciplines and with varying knowledge levels to participate in assessment and decision making on the same basis.
- Through their visual platform, **BNs also offer a transparent and justifiable evidence base** to inform management options and support practitioners' decision making processes.

The following potential applications of the models developed in the study demonstrate their utility for practitioners such as water utilities managers, regulatory authorities, treatment plant operators, risk modellers and public facility managers:

- Hazardous event conditions can be simulated in the model to determine likely risk outcomes;
- Treatment conditions can be modelled in conjunction with various exposure profiles to determine the type of public open space for which the water is most suited, for irrigation purposes;
- A desired risk outcome can be set along with other model constraints, to determine the conditions required in the remaining nodes to achieve the desired risk outcome;
- Sensitivity analysis can determine which step in the treatment chain has the most impact on the pathogen concentration at the end of treatment;
- Similarly, sensitivity analysis can determine which onsite risk reduction measure is most efficient, or rank the treatment chain steps and/or risk reduction measures in order of efficiency;
- Sensitivity analysis can indicate which variables in the model are optimal for further research expenditure in terms of improving risk estimates.

In summary, the research described in this thesis has fulfilled multiple modelling objectives: incorporation of several key exposure variables, use of the multiple barrier approach to risk management, adaptive management capability and integration of water treatment performance indicators with health-based targets. The explicit portrayal and quantification of complex exposure-health scenarios has improved understanding of public health risks associated with wastewater reuse scenarios. By offering a clearer understanding of role of systems/Bayesian approaches in characterising environmental exposures and public health risks, the thesis improves accessibility to the BN methodology, while increasing the clarity and accessibility of QMRA steps via the conceptual models. Use of secondary data to quantify the models in this research has the benefit of being low in cost and providing rapid results. The sensitivity analysis capabilities of BNs facilitate iterative model refinement, reflecting the iterative cycle of adaptive environmental management. The flexibility of BNs increases the scope of QMRA, for instance through the updating of poor quality priors by means of backwards inference. The flexibility of BN modelling in this domain is further established in the discussion of numerous alternative quantifying approaches possible with BNs, such as use of sewage treatment plant operational data, laboratory or field experimental data, expert opinions or hydraulic modelling data. This thesis establishes QMRA-based BNs as an accessible, transparent tool to facilitate 'fit-for-purpose' water recycling and offers a transparent, defensible evidence base for management options and decision making in regard to water recycling.

7.1 FURTHER RESEARCH

The body of literature on microbiological risk in wastewater reuse is extensive, however many aspects require further investigation. Moreover, as indicated earlier in Chapter 3 (Beaudequin et al., 2015a), applications of BNs in this domain are relatively novel and the field is somewhat open to exploration. In particular however, the following indications for further research have arisen in the course of this work and these are now briefly discussed in terms of expanding the methods and broadening the applications of the research.

A potential area for future research is a further series of stakeholder meetings to progress refinement and validation of the prototype BNs. The scope of the research, illustrated in Figure 1.1 in the Introduction, includes five of the six steps of

the contemporary risk assessment and management process (NRC, 2009). The sixth step of the process, risk communication, was beyond the scope of the work. Nevertheless, stakeholder and community engagement is a critical feature of risk communication, without which assessment findings and management decisions cannot be successfully translated to practice. Participatory refinement is a crucial component of the knowledge engineering cycle underpinning the BN methodology. If structured appropriately, subjective and qualitative information from stakeholders and experts can be used as additional inputs to the BNs, complementing the method used for BN quantification in this study. A second potential variation to the method would be extension of the networks to incorporate feedback loops, with the use of dynamic BNs. This would enable variable outputs to be used as eventual inputs where appropriate and is an important capability in modelling environmental systems.

A further application of this work is the use of BNs for characterisation of additional exposure pathways, both for wastewater uses other than irrigation and for routes of exposure other than ingestion. In addition to irrigation, nonpotable wastewater uses such as industrial cooling, aquaculture and fire protection are important applications for this important resource, for which further research on health risks associated with exposure is indicated. In addition to ingestion, gaps exist in understanding alternate wastewater exposure routes such as inhalation, inhalation-ingestion of aerosols and inadvertent ingestion of wastewater contaminated soils, either directly or through contact with surfaces or fomites.

A second future application of the research relates to the dose response step of a QMRA. A recurring finding in the present work was the ability and flexibility of BNs in the characterisation of exposure; the ability to represent the many influences on pathogen concentration and the various routes, frequencies, consumption quantities and other variables governing the final pathogen dose. This capacity of BNs could conceivably be extended to characterisation of influences such as immunocompetence or nutritional status on dose response in the individual, or characterisation of dose response in vulnerable groups, in the interests of more accurate and more credible health risk estimates. Although such data do not exist and/or are expensive, difficult or unethical to obtain, these variables could be parameterised through structured, formal elicitation from expert teams.

A third future application of the BNs developed in this study is the concomitant consideration of risk associated with chemical and microbiological contaminants, including their interactions. The simultaneous effects of wastewater chemical contaminants and microbes on human health, discussed earlier could be modelled using expert opinions on possible interactions. The development of antibiotic resistance in microbial communities of pathogens, opportunistic pathogens and environmental bacteria in wastewater is another facet of concern requiring investigation (Bouki et al., 2013, Gatica and Cytryn, 2013, Varela and Manaia, 2013). The chemical-microbiological interface in the wastewater domain is clearly an important area for further exploration.

Finally, a possible future application of the research is the embedding of a suite of further refined BNs in a probability-based decision support system with a graphic user interface. Such a tool might be web-based, with subscriptions available for purchase by councils, water utilities or other regulatory authorities, for the assessment and management of water recycling schemes. The decision support system could provide rapid scenario assessments, facilitating fit-for-purpose water recycling options.

Bibliography

ABBAS, H., NASR, R. & SEIF, H. 2006. Study of waste stabilization pond geometry for the wastewater treatment efficiency. *Ecological Engineering,* 28**,** 25-34.

AGUILERA, P. A., FERNÁNDEZ, A., FERNÁNDEZ, R., RUMÍ, R. & SALMERÓN, A. 2011. Bayesian networks in environmental modelling. *Environmental Modelling and Software,* 26**,** 1376-1388.

ALBERT, I., GRENIER, E., DENIS, J. B. & ROUSSEAU, J. 2008. Quantitative risk assessment from farm to fork and beyond: A global Bayesian approach concerning food-borne diseases. *Risk Analysis,* 28**,** 557-571.

ALCALDE, L., FOLCH, M. & TAPIAS, J. C. 2012. Removal and relationships of microbial indicators in a water treatment and reclamation facility. *Journal of Water and Health,* 10**,** 549-556.

ALSTON, C., MENGERSEN, K. L. & PETTITT, A. 2012. *Case studies in Bayesian statistical modelling and analysis,* Chichester, West Sussex, John Wiley & Sons.

ANASTAS, P., TEICHMAN, K. & COHEN HUBAL, E. 2010. Ensuring the safety of chemicals. *Journal of Exposure Science and Environmental Epidemiology,* 20**,** 395-396.

ANDERSON, R. M. & MAY, R. 1991. *Infectious Diseases of Humans: Dynamics and Control,* New York, Oxford University Press.

ASANO, T., BURTON, F., LEVERENZ, H. & METCALF & EDDY, I. 2007. *Water reuse: Issues, technologies, and applications,* New York, McGraw-Hill Professional, Access Engineering.

ASANO, T., LEONG, L. Y. C., RIGBY, M. G. & SAKAJI, R. H. 1992. Evaluation of the California wastewater reclamation criteria using enteric virus monitoring data. *Water Science and Technology,* 26**,** 1513-1524.

ASANO, T. & TCHOBANOGLOUS, G. 1991. The role of wastewater reclamation and reuse in the USA. *Water Science and Technology,* 23**,** 2049-2059.

ASHBOLT, N., SCHOEN, M., SOLLER, J. & ROSER, D. 2010. Predicting pathogen risks to aid beach management: The real value of quantitative microbial risk assessment (QMRA). *Water Research,* 44**,** 4692-4703.

ATMAR, R. 2010. Noroviruses: State of the art. *Food and Environmental Virology,* 2**,** 117-126.

ATMAR, R. L., OPEKUN, A. R., GILGER, M. A., ESTES, M. K., CRAWFORD, S. E., NEILL, F. H., RAMANI, S., HILL, H., FERREIRA, J. & GRAHAM, D. Y. 2014. Determination of the 50% human infectious dose for Norwalk virus. *Journal of Infectious Diseases,* 209**,** 1016-1022.

AUYANG, S. 2004. *Systems approach* [Online]. Available: http://www.creatingtechnology.org/sysapp.htm [Accessed 21 June 1014.

BAE, J. Y., LEE, J. S., SHIN, M. H., LEE, S. H. & HWANG, I. G. 2011. Effect of wash treatments on reducing human norovirus on iceberg lettuce and perilla leaf. *Journal of Food Protection,* 74**,** 1908-1911.

BARKER, G. & GOMEZ-TOME, N. 2013. A risk-assessment model for enterotoxigenic *Staphylococcus aureus* in pasteurized milk: a potential route to source-level inference. *Risk Analysis,* 33**,** 249-269.

BARKER, G. C. 2004. Application of Bayesian belief network models to food safety science. *In:* VAN BOEKEL, M., STEIN, A. & VAN BRUGGEN, H. (eds.) *Bayesian statistics and quality modelling in the agri-food production chain.* Dordrecht, The Netherlands: Kluwer.

BARKER, G. C., GOMEZ, N. & SMID, J. 2009. An introduction to biotracing in food chain systems. *Trends in Food Science and Technology,* 20**,** 220-226.

BARKER, G. C., TALBOT, N. L. C. & PECK, M. W. 2002. Risk assessment for *Clostridium botulinum:* a network approach. *International Journal of Biodeterioration and Biodegradation,* 50**,** 167-175.

BARKER, S. F. 2014a. Risk of norovirus gastroenteritis from consumption of vegetables irrigated with highly treated municipal wastewater-evaluation of methods to estimate sewage quality. *Risk Analysis,* 34**,** 803-817.

BARKER, S. F. 2014b. *Wastewater reuse: Modelling the human health risks associated with reuse applications.* Doctor of Philosophy, University of Melbourne.

BARKER, S. F., AMOAH, P. & DRECHSEL, P. 2014. A probabilistic model of gastroenteritis risks associated with consumption of street food salads in Kumasi, Ghana: Evaluation of methods to estimate pathogen dose from water, produce or food quality. *Science of the Total Environment,* 487**,** 130-142.

BARKER, S. F., O'TOOLE, J., SINCLAIR, M. I., LEDER, K., MALAWARAARACHCHI, M. & HAMILTON, A. J. 2013a. A probabilistic model of norovirus disease burden associated with greywater irrigation of home-produced lettuce in Melbourne, Australia. *Water Research,* 47**,** 1421-1432.

BARKER, S. F., PACKER, M., SCALES, P. J., GRAY, S., SNAPE, I. & HAMILTON, A. J. 2013b. Pathogen reduction requirements for direct potable reuse in Antarctica: Evaluating human health risks in small communities. *Science of the Total Environment,* 461-462**,** 723-733.

BARTON, D. N., KUIKKA, S., VARIS, O., UUSITALO, L., HENRIKSEN, H. J., BORSUK, M., DE LA HERA, A., FARMANI, R., JOHNSON, S. & LINNELL, J. D. 2012. Bayesian networks in environmental and resource management. *Integrated Environmental Assessment and Management,* 8**,** 418-429.

BARTRAND, T. A., ROSEN, J., MIEOG, J. & HARGY, T. Characterizing influent water quality for a water reuse operation. 2013 Water Quality Technology Conference and Exposition, WQTC 2013, 2013.

BASTOS, R. K. X., BEVILACQUA, P. D., SILVA, C. A. B. & SILVA, C. V. 2008. Wastewater irrigation of salad crops: Further evidence for the evaluation of the WHO guidelines. *Water Science and Technology,* 57**,** 1213-1219.

BEAUDEQUIN, D., HARDEN, F., ROIKO, A. & MENGERSEN, K. 2016. Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water. *Science of the Total Environment,* 541**,** 1393–1409.

BEAUDEQUIN, D., HARDEN, F., ROIKO, A., STRATTON, H., LEMCKERT, C. & MENGERSEN, K. 2015a. Beyond QMRA: Modelling microbial health risk as a complex system using Bayesian networks. *Environment International,* 80**,** 8-18.

BEAUDEQUIN, D., HARDEN, F., ROIKO, A., STRATTON, H., LEMCKERT, C. & MENGERSEN, K. 2015b. Modelling microbial health risk of wastewater reuse: A systems perspective. *Environment International,* 84**,** 131-141.

BEN-GAL, I. 2007. Bayesian networks. *In:* RUGGERI, F., FALTIN, F. & KENETT, R. (eds.) *Encyclopedia of statistics in quality and reliability.* Wiley & Sons.

BICHAI, F. & SMEETS, P. 2013. Using QMRA-based regulation as a water quality management tool in the water security challenge: Experience from the Netherlands and Australia. *Water Research,* 47**,** 7315-7326.

BITTON, G. 2005. Wastewater reuse. *Wastewater microbiology, Third Edition.* New Jersey: John Wiley & Sons, Inc.

BLACK, R. E., LEVINE, M. M., CLEMENTS, M. L., HUGHES, T. P. & BLASER, M. J. 1988. Experimental *Campylobacter jejuni* infection in humans. *Journal of Infectious Diseases,* 157**,** 472-479.

BOLTON, N. F., CROMAR, N. J., HALLSWORTH, P. & FALLOWFIELD, H. J. 2010. A review of the factors affecting sunlight inactivation of micro-organisms in waste stabilisation ponds: Preliminary results for enterococci. *Water Science and Technology,* 61**,** 885-890.

BOUKI, C., VENIERI, D. & DIAMADOPOULOS, E. 2013. Detection and fate of antibiotic resistant bacteria in wastewater treatment plants: A review. *Ecotoxicology and Environmental Safety,* 91**,** 1-9.

BOYLE, M., SICHEL, C., FERNÁNDEZ-IBÁÑEZ, P., ARIAS-QUIROZ, G. B., IRIARTE-PUÑA, M., MERCADO, A., UBOMBA-JASWA, E. & MCGUIGAN, K. G. 2008. Bactericidal effect of solar water disinfection under real sunlight conditions. *Applied and Environmental Microbiology,* 74**,** 2997-3001.

BRACHO, N., LLOYD, B. & ALDANA, G. 2006. Optimisation of hydraulic performance to maximise faecal coliform removal in maturation ponds. *Water Research,* 40**,** 1677-1685.

BRISSAUD, F., TOURNOUD, M. G., DRAKIDES, C. & LAZAROVA, V. 2003. Mixing and its impact on faecal coliform removal in a stabilisation pond. *Water Science and Technology,* 48**,** 75-80.

BROOKS, J. P., MCLAUGHLIN, M. R., GERBA, P. & PEPPER, I. L. 2012. Land application of manure and class B biosolids: An occupational and public quantitative microbial risk assessment. *Journal of Environmental Quality,* 41**,** 2009-2023.

BROWNLEE, J. 2007. The 'Pathogenic Exposure' paradigm. *Technical Report 070422A.* Melbourne, Australia: Swinburne University of Technology.

BUCHANAN, R. L., HAVELAAR, A. H., SMITH, M. A., WHITING, R. C. & JULIENS, E. 2009. The key events dose-response framework: Its potential for application to foodborne pathogenic microorganisms. *Critical Reviews in Food Science and Nutrition,* 49**,** 718-728.

BUCHANAN, R. L., SMITH, J. L. & LONG, W. 2000. Microbial risk assessment: Dose-response relations and risk characterization. *International Journal of Food Microbiology,* 58**,** 159-172.

BUREAU OF METEROLOGY. 2016. *Climate Resilient Water Sources* [Online]. Australia. Available: http://www.bom.gov.au/water/crews/ [Accessed 2 October 2016.

CAPRA, F., BUCKLEY, P. & BARLOW, Z. 2014. *Systems thinking* [Online]. Berkeley, California. Available: http://www.ecoliteracy.org/essays/systems-thinking [Accessed 13 November 2014.

CARLANDER, A., SCHÖNNING, C. & STENSTRÖM, T. A. 2009. Energy forest irrigated with wastewater: A comparative microbial risk assessment. *Journal of Water and Health,* 7**,** 413-433.

CENTER FOR ADVANCING MICROBIAL RISK ASSESSMENT. 2013a. *Center for Advancing Microbial Risk Assessment* [Online]. Michigan State University, East Lansing, Michigan. Available: http://www.camra.msu.edu/index.html [Accessed 20 January 2015.

CENTER FOR ADVANCING MICROBIAL RISK ASSESSMENT. 2013b. *QMRAwiki* [Online]. Available: http://qmrawiki.msu.edu/index.php?title=Quantitative_Microbial_Risk_Assessment [Accessed 15 October 2013.

CHEN, S. H. & POLLINO, C. A. 2012. Good practice in Bayesian network modelling. *Environmental Modelling & Software,* 37**,** 134-145.

CHERRIE, J. W., SEMPLE, S., CHRISTOPHER, Y., SALEEM, A., HUGHSON, G. & PHILIPS, A. 2006. How important is inadvertent ingestion of hazardous substances at work? *Annals of Occupational Hygiene,* 50**,** 693-704.

CODEX ALIMENTARIUS COMMISSION 1999. Principles and guidelines for the conduct of microbiological risk assessment. Rome, Italy: Food and Agriculture Organization and World Health Organization.

COHEN HUBAL, E. A., BARR, D. B., KOCH, H. M. & BAHADORI, T. 2011. The promise of exposure science. *Journal of Exposure Science and Environmental Epidemiology,* 21**,** 121-122.

COOK, A., DEVINE, B., RODRIGUEZ, C., ROSER, D., KHAN, S. & ASHBOLT, N. 2011. *Assessing the public health impacts of recycled water use. Final report,* Perth, University of Western Australia.

COOPER, R. C. & OLIVIERI, A. W. 1998. Infectious disease concerns in wastewater reuse. *In:* ASANO, T. (ed.) *Wastewater Reclamation and Reuse.* Boca Raton, Florida: CRC Press.

COUPÉ, V. M. H., VAN DER GAAG, L. C. & HABBERNA, J. D. F. 2000. Sensitivity analysis: An aid for belief-network quantification. *Knowledge Engineering Review,* 15**,** 215-232.

COVELLO, V. T. & MERKHOFER, M. W. 1993. *Risk assessment methods: Approaches for assessing health and environmental risks,* New York, Plenum Press.

COVEY, S. R. 1990. *The seven habits of highly effective people: restoring the character ethic,* Melbourne, Business Library.

COWLES, K., KASS, R. & O'HAGAN, T. 2009. *What is Bayesian analysis?* [Online]. International Society for Bayesian Analysis. Available: http://bayesian.org/Bayes-Explained [Accessed 17/10/13.

CRAGGS, R. J., ZWART, A., NAGELS, J. W. & DAVIES-COLLEY, R. J. 2004. Modelling sunlight disinfection in a high rate pond. *Ecological Engineering,* 22**,** 113-122.

CRAINICEANU, C. M., STEDINGER, J. R., RUPPERT, D. & BEHR, C. T. 2003. Modeling the U.S. national distribution of waterborne pathogen concentrations with application to *Cryptosporidium parvum*. *Water Resources Research,* 39**,** SWC21-SWC215.

CULLEN, A. C. & FREY, H. C. 1999. *Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs,* New York, Plenum Press.

CUNLIFFE, D. 2006. Managing health risks to consumers. *In:* STEVENS, D. (ed.) *Growing crops with reclaimed wastewater.* Collingwood, Victoria: CSIRO Publishing.

CURTIS, T. P., MARA, D. D. & SILVA, S. A. 1992a. The effect of sunlight on faecal coliforms in ponds: Implications for research and design. *Water Science and Technology,* 26**,** 1729-1738.

CURTIS, T. P., MARA, D. D. & SILVA, S. A. 1992b. Influence of pH, oxygen, and humic substances on ability of sunlight to damage fecal coliforms in waste stabilization pond water. *Applied and Environmental Microbiology,* 58**,** 1335-1343.

DA SILVA, A. K., LE GUYADER, F. S., LE SAUX, J. C., POMMEPUY, M., MONTGOMERY, M. A. & ELIMELECH, M. 2008. Norovirus removal and particle association in a waste stabilization pond. *Environmental Science and Technology,* 42**,** 9151-9157.

DAVIES-COLLEY, R. J., DONNISON, A. M. & SPEED, D. J. 2000. Towards a mechanistic understanding of pond disinfection. *Water Science and Technology,* 42**,** 149-158.

DAVIES-COLLEY, R. J., DONNISON, A. M., SPEED, D. J., ROSS, C. M. & NAGELS, J. W. 1999. Inactivation of faecal indicator micro-organisms in waste stabilisation ponds: Interactions of environmental factors with sunlight. *Water Research* 33**,** 1220-1230.

DAWSON, D. J., PAISH, A., STAFFELL, L. M., SEYMOUR, I. J. & APPLETON, H. 2005. Survival of viruses on fresh produce, using MS2 as a surrogate for norovirus. *Journal of Applied Microbiology,* 98**,** 203-209.

DE KEUCKELAERE, A., JACXSENS, L., AMOAH, P., MEDEMA, G., MCCLURE, P. & JAYKUS, L. 2015. Zero risk does not exist: Lessons learned from microbial risk assessment related to use of water and safety of fresh produce. *Comprehensive Reviews in Food Science and Food Safety,* 14**,** 387-410.

DEERE, D., STEVENS, M., DAVISON, A., HELM, G. & DUFOUR, A. 2001. Management strategies. *In:* FEWTRELL, L. & BARTRAM, J. (eds.) *Water quality: guidelines, standards and health.* World Health Organisation.

DELIGNETTE-MULLER, M. L. & CORNU, M. 2008. Quantitative risk assessment for *Escherichia coli* O157:H7 in frozen ground beef patties consumed by young children in French households. *International Journal of Food Microbiology,* 128**,** 158-164.

DELIGNETTE-MULLER, M. L., CORNU, M., POUILLOT, R. & DENIS, J. B. 2006. Use of Bayesian modelling in risk assessment: Application to growth of *Listeria monocytogenes* and food flora in cold-smoked salmon. *International Journal of Food Microbiology,* 106**,** 195-208.

DONALD, M., COOK, A. & MENGERSEN, K. 2009. Bayesian network for risk of diarrhea associated with the use of recycled water. *Risk Analysis,* 29**,** 1672-1685.

DRECHSEL, P., QADIR, M. & WICHELNS, D. 2015. Wastewater: Economic asset in an urbanizing world. The Netherlands: Springer.

DRECHSEL, P., SCOTT, C. A., RASCHID-SALLY, L., REDWOOD, M. & BAHRI, A. 2010. *Wastewater irrigation and health. Assessing and mitigating risk in low-income countries,* Ottowa, Canada, International Water Management Institute and International Development Research Centre.

DUDLEY, R. H., HEKIMIAN, K. K. & MECHALAS, B. J. 1976. A scientific basis for determining recreational water quality criteria. *Journal of Water Pollution Control Federation,* 48**,** 2661-2677.

DUFFY, G., BUTLER, F. & CUMMINS, E. 2006. *E. coli* O157:H7 in beefburgers produced in the Republic of Ireland: a quantitative microbial risk assessment. Dublin, Ireland.: Ashtown Food Research Centre, Teagasc, The Agriculture and Food Development Authority.

DÜSPOHL, M., FRANK, S. & DÖLL, P. 2012. A review of Bayesian networks as a participatory modeling approach in support of sustainable environmental management. *Journal of Sustainable Development,* 5**,** 1-18.

EISENBERG, J. N., BROOKHART, M. A., RICE, G., BROWN, M. & COLFORD, J. M. 2002. Disease transmission models for public health decision making: Analysis of epidemic and endemic conditions caused by waterborne pathogens. *Environmental Health Perspectives,* 110**,** 783-790.

EISENBERG, J. N., LEI, X., HUBBARD, A. H., BROOKHART, M. A. & COLFORD, J. M. 2005. The role of disease transmission and conferred immunity in outbreaks: Analysis of the 1993 *Cryptosporidium* outbreak in Milwaukee, Wisconsin. *American Journal of Epidemiology,* 161.

ELLISON 2004. Bayesian inference in ecology. *Ecology Letters, (2004) 7: 509–520*.

ENHEALTH 2012. Environmental health risk assessment: Guidelines for assessing human health risk from environmental hazards. Commonwealth of Australia. Department of Health and Ageing.

ENVIRONMENTAL PROTECTION AGENCY QUEENSLAND 2005. Queensland water recycling guidelines. Brisbane, Queensland: The State of Queensland. Environmental Protection Agency.

FENTON, N. & NEIL, M. 2013. *Risk assessment and decision analysis with Bayesian networks,* Boca Raton, Florida, CRC Press.

FERNANDEZ, A., TEJEDOR, C. & CHORDI, A. 1992. Effect of different factors on the die-off of fecal bacteria in a stabilization pond purification plant. *Water Research,* 26**,** 1093-1098.

FEWTRELL, L., BARTRAM, J. & WORLD HEALTH ORGANIZATION 2001. *Water quality: Guidelines, standards, and health - risk assessment and management for water-related infectious disease,* London, United Kingdom, IWA Publishing.

FOOD AND AGRICULTURE ORGANIZATION/WORLD HEALTH ORGANIZATION 2003. Hazard characterization for pathogens in food and water: Guidelines. *Microbiological Risk Assessment Series No. 3*.

FOOD AND AGRICULTURE ORGANIZATION/WORLD HEALTH ORGANIZATION 2008. Exposure assessment of microbiological hazards in food: Guidelines. *Microbiological Risk Assessment Series No. 7*.

FUHS, O. W. 1975. A probabilistic model of bathing beach safety. *Science of the Total Environment,* 4**,** 165-175.

GARCIA, S., LUENGO, J., SAEZ, J. A. L., V. & HERRERA, F. 2013. A survey of discretization techniques: Taxonomy and empirical analysis in supervised learning. *IEEE Transactions on Knowledge and Data Engineering,* 25**,** 734-750.

GARCIA, X. & PARGAMENT, D. 2015. Reusing wastewater to cope with water scarcity: Economic, social and environmental considerations for decision-making. *Resources, Conservation and Recycling,* 101**,** 154-166.

GATICA, J. & CYTRYN, E. 2013. Impact of treated wastewater irrigation on antibiotic resistance in the soil microbiome. *Environmental Science and Pollution Research,* 20**,** 3529-3538.

GENTHE, B., LE ROUX, W. J., SCHACHTSCHNEIDER, K., OBERHOLSTER, P. J., ANECK-HAHN, N. H. & CHAMIER, J. 2013. Health risk implications from simultaneous exposure to multiple environmental contaminants. *Ecotoxicology and Environmental Safety,* 93**,** 171-179.

GERBA, C. P. 6 November 2015. *RE: Concurrent microbial exposures.* Type to BEAUDEQUIN, D.

GERBA, C. P., KITAJIMA, M. & IKER, B. 2013. Viral presence in waste water and sewage and control methods. *Viruses in food and water: Risks, surveillance and control.*

GERBA, C. P., ROSE, J. B. & HAAS, C. N. 1996. Sensitive populations: who is at the greatest risk? *International Journal of Food Microbiology,* 30**,** 113-123.

GLOYNA, E. 1971. Waste stabilization ponds. *World Health Organization,* Monograph Series No. 60.

GORMAN, N. M., VAN TONGEREN, M. & SEMPLE, S. 2014. Simulated transfer of liquids and powders from hands and clothing to the mouth. *Journal of Occupational and Environmental Hygiene,* 11**,** 633-644.

GOULDING, R., JAYASURIYA, N. & HORAN, E. 2012. A Bayesian network model to assess the public health risk associated with wet weather sewer overflows discharging into waterways. *Water Research,* 46**,** 4933-4940.

GREINER, M., SMID, J., HAVELAAR, A. H. & MÜLLER-GRAF, C. 2013. Graphical models and Bayesian domains in risk modelling: Application in microbiological risk assessment. *Preventive Veterinary Medicine,* 110**,** 4-11.

GRONEWOLD, A. D., BORSUK, M. E., WOLPERT, R. L. & RECKHOW, K. H. 2008. An assessment of fecal indicator bacteria-based water quality standards. *Environmental Science and Technology,* 42**,** 4676-4682.

GRONEWOLD, A. D., MYERS, L., SWALL, J. L. & NOBLE, R. T. 2011. Addressing uncertainty in fecal indicator bacteria dark inactivation rates. *Water Research,* 45**,** 652-664.

GRONEWOLD, A. D., STOW, C. A., VIJAYAVEL, K., MOYNIHAN, M. A. & KASHIAN, D. R. 2013. Differentiating *Enterococcus* concentration spatial, temporal, and analytical variability in recreational waters. *Water Research,* 47**,** 2141-2152.

HAAS, C. N. 1996. Acceptable microbial risk. *Journal of the American Water Works Association,* 88**,** 8.

HAAS, C. N. 2002. Progress and data gaps in quantitative microbial risk assessment. *Water Science and Technology,* 46**,** 277.

HAAS, C. N. 2011. *Basic microbial dose response [slides],* Drexel University, Philadelphia, U.S.A.

HAAS, C. N. 5 November 2015. *RE: Concurrent microbial exposures.* Type to BEAUDEQUIN, D.

HAAS, C. N., ROSE, J. B. & GERBA, C. P. 1999. *Quantitative microbial risk assessment,* New York, John Wiley and Sons Inc.

HAAS, C. N., ROSE, J. B. & GERBA, C. P. 2014. *Quantitative microbial risk assessment,* New Jersey, John Wiley & Sons, Inc.

HALL, G., KIRK, M. D., BECKER, B. N., GREGORY, J. E., UNICOMB, L., MILLARD, G., STAFFORD, R., LALOR, K. & GROUP, O. 2005. Estimating foodborne gastroenteritis, Australia. *Emerging Infectious Diseases,* 11**,** 1257-64.

HAMER, W. 1906. Variability and persistence of type in epidemic diseases: Measles in London. Milroy Lecture, Royal College of Physicians, London, Bedford Press.

HAMILTON, A. J., STAGNITTI, F., KUMARAGE, S. C. & PREMIER, R. R. 2007. RIRA: A tool for conducting health risk assessments for irrigation of edible crops with recycled water. *Computers and Electronics in Agriculture,* 57**,** 80-87.

HAMILTON, A. J., STAGNITTI, F., PREMIER, R., BOLAND, A. M. & HALE, G. 2006. Quantitative microbial risk assessment models for consumption of raw vegetables irrigated with reclaimed water. *Applied and Environmental Microbiology,* 72**,** 3284-3290.

HARWOOD, V. J., LEVINE, A. D., SCOTT, T. M., CHIVUKULA, V., LUKASIK, J., FARRAH, S. R. & ROSE, J. B. 2005. Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Applied and Environmental Microbiology,* 71**,** 3163-3170.

HAVELAAR, A. H. 2012. *QMRA - a framework for assessing microbiological public health risks,* Dutch National Institute for Public Health and the Environment (RIVM) and Utrecht University, The European College of Veterinary Public Health Annual Scientific Conference, Maastricht, 23 August 2012.

HAVELAAR, A. H., EVERSA, E. G. & NAUTA, M. J. 2008. Challenges of quantitative microbial risk assessment at EU level. *Trends in Food Science and Technology,* 19**,** S26-S33.

HAVELAAR, A. H. & MELSE, J. M. 2003. Quantifying public health risk in the WHO guidelines for drinking-water quality: A burden of disease approach *RIVM report 734301022.* The Netherlands: Rijksinstituut voor Volksgezondheid en Milieu, Netherlands.

HAVELAAR, A. H. & SWART, A. N. 2014. Impact of acquired immunity and dose-dependent probability of illness on quantitative microbial risk assessment. *Risk Analysis,* 34**,** 1807–1819.

HENRIKSEN, H. J., ZORRILLA-MIRAS, P., DE LA HERA, A. & BRUGNACH, M. 2012. Use of Bayesian belief networks for dealing with ambiguity in integrated groundwater management. *Integrated Environmental Assessment and Management,* 8**,** 430-444.

HETHCOTE, H. W. 1976. Quantitative analyses of communicable disease models. *Mathematical Biosciences,* 28**,** 335-356.

HETHCOTE, H. W. 2000. The mathematics of infectious diseases. *Siam Review,* 42**,** 599-653.

HOLLING, C. S. 1978. *Adaptive environmental assessment and management,* New York, Wiley.

HUNTER, P. R., DE SYLOR, M. A., RISEBRO, H. L., NICHOLS, G. L., KAY, D. & HARTEMANN, P. 2011. Quantitative microbial risk assessment of cryptosporidiosis and giardiasis from very small private water supplies. *Risk Analysis,* 31**,** 228-236.

HUNTER, P. R. & FEWTRELL, L. 2001. Acceptable risk. *Water quality: Guidelines, standards and health.* World Health Organisation.

HUTCHISON, M. L., WALTERS, L. D., MOORE, T., THOMAS, D. J. I. & AVERY, S. M. 2005. Fate of pathogens present in livestock wastes spread onto fescue plots. *Applied and Environmental Microbiology,* 71**,** 691-696.

HUXEDURP, L. M., PALSDOTTIR, G. & ALTAVILLA, N. 2014. Risk-based planning for water recycling in an Australian context. *Water Science and Technology: Water Supply,* 14**,** 971-983.

INSTITUTE OF MEDICINE 2013. *Environmental decisions in the face of uncertainty,* Washington, DC, National Academies Press.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE 2014. Summary for policy makers. *Climate change 2014: Impacts, adaptation and vulnerability. Part A: Global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report***,** 1-32.

INTERNATIONAL LIFE SCIENCE INSTITUTE 2000. Revised framework for microbial risk assessment. *An ILSI Workshop Report.* Washington, DC.

JACKSON, E. F. & JACKSON, C. R. 2008. Viruses in wetland ecosystems. *Freshwater Biology,* 53.

JAKEMAN, A. J., LETCHER, R. A. & CHEN, S. 2007. Integrated assessment of impacts of policy and water allocation change across social, economic and environmental dimensions. *In:* HUSSEY, K. & DOVERS, S. (eds.) *Managing water for Australia: The social and institutional challenges.* Australian National University, Canberra: CSIRO Publishing.

JAKEMAN, A. J., LETCHER, R. A. & NORTON, J. P. 2006. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling and Software,* 21**,** 602-614.

JENKINS, M. B., LIOTTA, J. L. & BOWMAN, D. D. 2013. Inactivation kinetics of *Cryptosporidium parvum* oocysts in a swine waste lagoon and spray field. *Journal of Parasitology,* 99**,** 337-342.

JENSEN, F. 2001. *Bayesian networks and decision graphs,* New York, Springer-Verlag.

JENSEN, F. V. & NIELSEN, T. D. 2007. *Bayesian networks and decision graphs,* New York, Springer.

JOHNSON, S., FIELDING, F., HAMILTON, G. S. & MENGERSEN, K. 2010. An integrated Bayesian network approach to *Lyngbya majuscule* bloom initiation. *Marine Environmental Research,* 69**,** 27-37.

JOHNSON, S., LOW-CHOY, S. & MENGERSEN, K. 2012. Integrating Bayesian networks and geographic information systems: Good practice examples. *Integrated Environmental Assessment and Management,* 8**,** 473-479.

JOHNSON, S. & MENGERSEN, K. 2011. Integrated Bayesian network framework for modelling complex ecological issues. *Integrated Environmental Assessment and Management,* 8**,** 480-490.

JORI, G. & BROWN, S. B. 2004. Photosensitized inactivation of microorganisms. *Photochemical and Photobiological Sciences,* 3**,** 403-405.

JULIENS, E., BOOBIS, A. R. & OLIN, S. S. 2009. The key events dose-response framework: A cross-disciplinary mode-of-action based approach to examining dose-response and thresholds. *Critical Reviews in Food Science and Nutrition,* 49**,** 682-689.

KADIR, K. & NELSON, K. L. 2014. Sunlight mediated inactivation mechanisms of *Enterococcus faecalis* and *Escherichia coli* in clear water versus waste stabilization pond water. *Water Research,* 50**,** 307-317.

KEITHLIN, J., SARGEANT, J., THOMAS, M. K. & FAZIL, A. 2014. Chronic sequelae of *E. coli* O157: Systematic review and meta-analysis of the proportion of *E. coli* O157 cases that develop chronic sequelae. *Foodborne Pathogens and Disease,* 11**,** 79-95.

KERAITA, B., DRECHSEL, P. & KONRADSEN, F. 2010. Up and down the sanitation ladder: Harmonizing the treatment and multiple-barrier perspectives on risk reduction in wastewater irrigated agriculture. *Irrigation and Drainage Systems,* 24**,** 23-35.

KITAJIMA, M., IKER, B. C., PEPPER, I. L. & GERBA, C. P. 2014. Relative abundance and treatment reduction of viruses during wastewater treatment processes - Identification of potential viral indicators. *Science of the Total Environment,* 488-489**,** 290-296.

KJAERULFF, U. & VAN DER GAAG, L. C. Making sensitivity analysis computationally efficient. Sixteenth Conference on Uncertainty in Artificial Intelligence 2000, 2000 San Francisco, California.

KONTKANEN, P., MYLLYMAKI, P., SILANDER, T. & TIRRI, H. Comparing predictive inference methods for discrete domains. Sixth International Workshop on Artificial Intelligence and Statistics, 1997 Ft. Lauderdale, USA. 311-318.

KOOPMAN, J. S., CHICK, S. E., SIMON, C. P., RIOLO, C. S. & JACQUEZ, G. 2002. Stochastic effects on endemic infection levels of disseminating versus local contacts. *Mathematical Biosciences,* 180**,** 49-71.

KORB, K. B. & NICHOLSON, A. E. 2011. *Bayesian artificial intelligence,* Boca Raton, Florida, CRC Press.

KUIKKA, S., HILDEN, M., GISLASON, H., HANSSON, S., SPARHOLT, H. & VARIS, O. 1999. Modeling environmentally driven uncertainties in Baltic cod (*Gadus morhua*) management by Bayesian influence diagrams. *Canadian Journal of Fisheries and Aquatic Sciences,* 56**,** 629-641.

LAMMERDING, A. M. 1997. An overview of microbial food safety risk assessment. *Journal of Food Protection,* 60**,** 1420-1425.

LANDUYT, D., BOREKX, S., D'HONDT, R., ENGELEN, G., AERTSENS, J. & GOETHALS, P. 2013. A review of Bayesian belief networks in ecosystem service modelling. *Environmental Modelling and Software,* 46**,** 1-11.

LANIAK, G. F., OLCHIN, G., GOODALL, J., VOINOV, A., HILL, M., GLYNN, P., WHELAN, G., GELLER, G., QUINN, N. W. T., BLIND, M., PECKHAM, S., REANEY, S., GABER, N., KENNEDY, R. & HUGES, A. 2013. Integrated environmental modeling: A vision and roadmap for the future. *Environmental Modelling and Software,* 39**,** 3-23.

LECHEVALLIER, M. & BUCKLEY, M. 2007. *Clean water - what is acceptable microbial risk?,* Washington, D. C., American Academy of Microbiology.

LEMPERT, R. J. & COLLINS, M. T. 2007. Managing the risk of uncertain threshold responses: Comparison of robust, optimum, and precautionary approaches. *Risk Analysis,* 27**,** 1009-1026.

LI, D., YANG, H. & LIANG, X. 2010. Diagnosis and prognostic of wastewater treatment system based on Bayesian network. *AIP Conference Proceedings,* 1251**,** 25-28.

LI, D., YANG, H. Z. & LIANG, X. F. 2013. Prediction analysis of a wastewater treatment system using a Bayesian network. *Environmental Modelling and Software,* 40**,** 140-150.

LINDSAY, J. A. 1997. Chronic sequelae of foodborne disease. *Emerging Infectious Diseases,* 3**,** 443-452.

LIU, L., GAO, Y. & WANG, Y. 2013. Improvement and application of modular process risk modeling method for microbial risk assessment. *Journal of Chemical and Pharmaceutical Research,* 5**,** 434-438.

LLOYD, B., VORKAS, C. & GUGANESHARAJAH, K. 2003. Reducing hydraulic short-circuiting in maturation ponds to maximize pathogen removal using channels and wind breaks. *Water Science and Technology,* 48**,** 153-162.

LOW CHOY, S., O'LEARY, R. & MENGERSEN, K. 2009. Elicitation by design in ecology: Using expert opinion to inform priors for Bayesian statistical models. *Ecology,* 90**,** 265-277.

MAÏGA, Y., DENYIGBA, K., WETHE, J. & OUATTARA, A. S. 2009a. Sunlight inactivation of *Escherichia coli* in waste stabilization microcosms in a sahelian region (Ouagadougou, Burkina Faso). *Journal of Photochemistry and Photobiology B: Biology,* 94**,** 113-119.

MAÏGA, Y., WETHE, J., DENYIGBA, K. & OUATTARA, A. S. 2009b. The impact of pond depth and environmental conditions on sunlight inactivation of *Escherichia coli* and enterococci in wastewater in a warm climate. *Canadian Journal of Microbiology,* 55**,** 1364-1374.

MARA, D. 2009. Waste stabilization ponds: Past, present and future. *Desalination and Water Treatment,* 4**,** 85-88.

MARA, D. 2011. Water and wastewater-related disease and infection risks: What is an appropriate value for the maximum tolerable additional burden of disease? *Journal Of Water and Health,* 9**,** 217-224.

MARA, D. & SLEIGH, A. 2010. Estimation of norovirus infection risks to consumers of wastewater-irrigated food crops eaten raw. *Journal of Water and Health,* 8**,** 39-43.

MARA, D. D. 2003. Low-cost treatment systems. *In:* MARA, D. & HORAN, N. (eds.) *The handbook of water and wastewater microbiology.* London: Elsevier, Academic Press.

MARA, D. D. 2004. *Domestic wastewater treatment in developing countries,* London, Earthscan.

MARA, D. D. & FEACHEM, R. G. 2003. Unitary environmental classification of water-and excreta-related communicable diseases. *In:* MARA, D. & HORAN, N. (eds.) *The Handbook of Water and Wastewater Microbiology.* London: Elsevier, Academic Press.

MARA, D. D., SLEIGH, P. A., BLUMENTHAL, U. J. & CARR, R. M. 2007. Health risks in wastewater irrigation: Comparing estimates from quantitative microbial risk analyses and epidemiological studies. *Journal of Water and Health,* 5**,** 39-50.

MARCOT, B. G., STEVENTON, J. D., SUTHERLAND, G. D. & MCANN, R. K. 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. *Canadian Journal of Forest Research,* 36**,** 3063-3074.

MCCANN, R. K., MARCOT, B. G. & ELLIS, R. 2006. Bayesian belief networks: applications in ecology and natural resource management. *Canadian Journal of Forest Research,* 36**,** 3053-3062.

MEADOWS, D. H. 2008. *Thinking in systems: A primer,* Vermont, U.S.A., Chelsea Green Publishing.

MEDEMA, G. J. & ASHBOLT, N. 2006. QMRA: Its value for risk management. *In:* MICRORISK (ed.). The Netherlands: Kiwa Water Research.

MESSNER, M. J., CHAPPELL, C. L. & OKHUYSEN, P. C. 2001. Risk assessment for *Cryptosporidium*: a hierarchical Bayesian analysis of human response data. *Water Research,* 35**,** 3934-3940.

MITCHELL-BLACKWOOD, J., GURIAN, P. L., LEE, R. & THRAN, B. 2012. Variance in *Bacillus anthracis* virulence assessed through Bayesian hierarchical dose-response modelling. *Journal of Applied Microbiology,* 113**,** 265-275.

MOK, H., BARKER, S. & HAMILTON, A. 2014. A probabilistic quantitative microbial risk assessment model of norovirus disease burden from wastewater irrigation of vegetables in Shepparton, Australia. *Water Research,* 54**,** 347-362.

MOK, H. & HAMILTON, A. 2014. Exposure factors for wastewater-irrigated Asian vegetables and a probabilistic rotavirus disease burden model for their consumption. *Risk Analysis,* 34**,** 602-613.

MORENO, M. D. 1990. A tracer study of the hydraulics of facultative stabilization ponds. *Water Research,* 24**,** 1025-1030.

MUÑOZ, I., TOMÀS, N., MAS, J., GARCÍA-REYES, J. F., MOLINA-DÍAZ, A. & FERNÁNDEZ-ALBA, A. R. 2010. Potential chemical and microbiological risks on human health from urban wastewater reuse in agriculture. Case study of wastewater effluents in Spain. *Journal of Environmental Science and Health. Part. B, Pesticides, food contaminants, and agricultural wastes,* 45**,** 300-309.

NATIONAL HEALTH AND MEDICAL RESEARCH COUNCIL 2004. Australian Drinking Water Guidelines *In:* NATIONAL HEALTH AND MEDICAL RESEARCH COUNCIL (ed.). Canberra, Australian Capital Territory: Government of Australia.

NATIONAL RESEARCH COUNCIL 1983. *Risk assessment in the federal government: Managing the process,* Washington, National Academies Press.

NATIONAL RESEARCH COUNCIL 1999. *Setting priorities for drinking water contaminants,* Washington, DC, National Academies Press.

NATIONAL RESEARCH COUNCIL 2004. *Indicators for waterborne pathogens,* Washington, National Academies Press.

NATIONAL RESEARCH COUNCIL 2009. *Science and decisions: Advancing risk assessment,* Washington DC, National Academies Press.

NATIONAL RESEARCH COUNCIL COMMITTEE ON INDICATORS FOR WATERBORNE PATHOGENS 2004. *Indicators for waterborne pathogens,* Washington, DC National Academies Press.

NATIONAL WATER COMMISSION 2014. National performance report 2012-13: urban water utilities. NWC, Canberra, Australia.

NAUTA, M. J. 2000. Separation of uncertainty and variability in quantitative microbial risk assessment models. *International Journal of Food Microbiology,* 57 9-18.

NAUTA, M. J. 2001. A modular process risk model structure for quantitative microbiological risk assessment and its application in an exposure assessment of *Bacillus cereus* in a REPFED. *Quantitative safety aspect of pathogens in foods.* Bilthoven, The Netherlands: National Institute of Public Health and the Environment.

NAUTA, M. J., JACOBS-REITSMA, W. F. & HAVELAAR, A. H. 2007. A risk assessment model for *Campylobacter* in broiler meat. *Risk Analysis,* 27**,** 845-861.

NEWTON, A. C. 2009. Bayesian belief networks in environmental modelling: a review of recent progress. *In:* FINDLEY, P. N. (ed.) *Environmental modelling: new research.* New York: Nova Science Publishers Inc.

NICHOLSON, A. E. & FLORES, M. J. 2011. Combining state and transition models with dynamic Bayesian networks. *Ecological Modelling,* 222**,** 555-566.

NRMMC-EPHC-AHMC 2006. *Australian guidelines for water recycling: Managing health and environmental risks (Phase 1),* Canberra, Australia, Natural Resource Management Ministerial Council Environment Protection and Heritage Council and Australian Health and Medical Council.

NWACHUKU, N. & GERBA, C. P. 2004. Microbial risk assessment: Don't forget the children. *Current Opinion in Microbiology,* 7**,** 206-209.

NYBERG, J. B., MARCOT, B. G. & SULYMA, R. 2006. Using Bayesian belief networks in adaptive management. *Canadian Journal of Forest Research,* 36**,** 3104-3104.

O'HAGAN, A., BUCK, C. E., DANESHKHAH, A., EISER, J. R., GARTHWAITE, P. H., JENKINSON, D. J., OAKLEY, J. E. & RAKOW, T. 2006. *Uncertain judgements: Eliciting experts' probabilities,* Chichester, England, Wiley.

O'TOOLE, J. 2011. *Identifying data gaps and refining estimates of pathogen health risks for alternative water sources,* Canberra, Waterlines Report, National Water Commission.

O'TOOLE, J., SINCLAIR, M., BARKER, F. & LEDER, K. 2014. Advice to risk assessors modeling viral health risk associated with household graywater. *Risk Analysis,* 34**,** 797-802.

O'TOOLE, J., SINCLAIR, M. & LEDER, K. 2008. Recycled water exposure: Filling the data gaps. *Water,* 35**,** 52-57.

OLUKANNI, D. O. & DUCOSTE, J. J. 2011. Optimization of waste stabilization pond design for developing nations using computational fluid dynamics. *Ecological Engineering,* 37**,** 1878-1888.

ONG, C. W. 2013. Norovirus: A challenging pathogen. *Healthcare Infection,* 18**,** 133-142.

PAGE, D., SIDHU, J. P. S. & TOZE, S. 2014. Microbial risk reduction of withholding periods during public open space irrigation with recycled water. *Urban Water Journal,* 12**,** 581-587.

PARSONS, D. J., ORTON, T. G., D'SOUZA, J., MOORE, A., JONES, R. & DODD, C. E. R. 2005. A comparison of three modelling approaches for quantitative risk assessment using the case study of *Salmonella* spp. in poultry meat. *International Journal of Food Microbiology,* 98**,** 35-51.

PATEL, M. M., HALL, A. J., VINJÉ, J. & PARASHAR, U. D. 2009. Noroviruses: A comprehensive review. *Journal of Clinical Virology,* 44**,** 1-8.

PAVIONE, D. M. S., BASTOS, R. K. X. & BEVILACQUA, P. D. 2013. Quantitative microbial risk assessment applied to irrigation of salad crops with waste stabilization pond effluents. *Water Science and Technology,* 67**,** 1208-1215.

PAYMENT, P. & LOCAS, A. 2011. Pathogens in water: Value and limits of correlation with microbial indicators. *Ground Water,* 49**,** 4-11.

PEARL, J. 1988. *Probabilistic reasoning in intelligent systems: Networks of plausible inference,* San Mateo, Calif, Morgan Kaufmann Publishers.

PEARL, J. 2000. *Causality,* Cambridge, United Kingdom, Cambridge University Press.

PEARSON, H. W., MARA, D. D., MILLS, S. W. & SMALLMAN, D. J. 1987. Physico-chemical parameters influencing faecal bacterial survival in waste stabilization ponds. *Water Science and Technology,* 19**,** 145-152.

PETTERSON, S. A. 2016. *RE: Concurrent microbial risk.* Type to ROIKO, A.

PETTERSON, S. A., SIGNOR, R., ASHBOLT, N. & ROSER, D. 2006. QMRA methodology. *Microrisk.* Sydney, Australia: University of New South Wales.

PETTERSON, S. R., ASHBOLT, N. J. & SHARMA, A. 2001. Microbial risks from wastewater irrigation of salad crops: A screening-level risk assessment. *Water Environment Research,* 73**,** 667-672.

PETTERSON, S. R., ASHBOLT, N. J. & SHARMA, A. 2002. Discussion of 'Microbial risks from wastewater irrigation of salad crops: A screening-level risk assessment'. *Water Environment Research,* 74**,** 411-411.

PETTERSON, S. R., DUMOUTIER, N., LORET, J. F. & ASHBOLT, N. J. 2009. Quantitative Bayesian predictions of source water concentration for QMRA from presence/absence data for *E. coli* O157:H7. *Water Science and Technology,* 59**,** 2245-2252.

PETTERSON, S. R., SIGNOR, R. S. & ASHBOLT, N. J. 2007. Incorporating method recovery uncertainties in stochastic estimates of raw water protozoan concentrations for QMRA. *Journal of Water and Health,* 5**,** 51-65.

PITCHFORTH, J. & MENGERSEN, K. 2013. A proposed validation framework for expert elicited Bayesian Networks. *Expert Systems with Applications,* 40**,** 162-167.

POLLINO, C. & HENDERSON, C. 2010. Bayesian networks: A guide for their application in natural resource management and policy. *Landscape Logic, Technical Report.* Hobart, Tasmania.

POLLINO, C. A. & HART, B. T. 2005. Bayesian decision networks - Going beyond expert elicitation for parameterisation and evaluation of ecological endpoints *In:* VOINOV, A., JAKEMAN, A. J., RIZZOLI, A. & SLEIGH, P. A. (eds.) *Third Biennial Meeting: Summit on Environmental Modelling and Software.* Burlington, USA: Water Studies Centre, Monash University, Clayton, Victoria.

POLLINO, C. A. & HART, B. T. 2008. Developing Bayesian network models within a risk assessment framework. *Fourth Biennial Meeting of International Environmental Modelling and Software Society.* Barcelona, Cantelonia.

POLLINO, C. A., WOODBERRY, O., NICHOLSON, A., KORB, K. & HART, B. T. 2007. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. *Environmental Modelling & Software,* 22**,** 1140-1152.

POSTEL, S. L. 2000. Entering an era of water scarcity: The challenges ahead. *Ecological Applications,* 10**,** 941–948.

POUILLOT, R., ALBERT, I., CORNU, M. & DENIS, J. B. 2003. Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to *Listeria monocytogenes*. *International Journal of Food Microbiology,* 81**,** 87-104.

POUILLOT, R. & DELIGNETTE-MULLER, M. L. 2010. Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages. *International Journal of Food Microbiology,* 142**,** 330-340.

PREDMORE, A. & LI, J. 2011. Enhanced removal of a human norovirus surrogate from fresh vegetables and fruits by a combination of surfactants and sanitzers. *Applied and Environmental Microbiology,* 77**,** 4829-4838.

PUJOL, J. M., EISENBERG, J. E., HAAS, C. N. & KOOPMAN, J. S. 2009. The effect of ongoing exposure dynamics in dose response relationships. *PLoS Computational Biology* [Online], 5. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-67650860432&partnerID=40&md5=ee7aabfd596f3ea062edfc79f9045f8f [Accessed 13 August 2013].

QIN, D., BLISS, P. J., BARNES, D. & FITZGERALD, P. A. 1991. Bacterial (total coliform) die off in maturation ponds. *Water Science and Technology,* 23**,** 1525-1534.

QUEENSLAND GOVERNMENT DEPARTMENT OF AGRICULTURE AND FISHERIES. 2010. Frequently asked questions about lettuce irrigation. Available: https://www.daf.qld.gov.au/plants/fruit-and-vegetables/vegetables/lettuce/faqs/irrigation [Accessed 9 April 2015].

RADCLIFFE, J. C. 2004. Water recycling in Australia. Parkville, Victoria: Australian Academy of Technical Sciences and Engineering.

RAJAN, S. R. & LETOURNEAU, D. K. 2012. What risk assessments of genetically modified organisms can learn from institutional analyses of public health risks. *Journal of Biomedicine and Biotechnology* [Online], 2012. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84870209814&partnerID=40&md5=f41d96124b9dbc5890c572807c4bf03b [Accessed 24 October 2015].

REINOSO, R., BLANCO, S., TORRES-VILLAMIZAR, L. A. & BÉCARES, E. 2011. Mechanisms for parasites removal in a waste stabilisation pond. *Microbial Ecology,* 61**,** 684-692.

RENOOIJ, S. 2001. Probability elicitation for belief networks: issues to consider. *The Knowledge Engineering Review,* 16**,** 255-269.

RIGAUX, C., ANCELET, S., CARLIN, F., NGUYEN-THÉ, C. & ALBERT, I. 2012a. Inferring an augmented Bayesian network to confront a complex quantitative microbial risk assessment model with durability studies: Application to *Bacillus cereus* on a courgette purée production chain. *Risk Analysis,* 33**,** 877-892.

RIGAUX, C., DENIS, J. B., ALBERT, I. & CARLIN, F. 2012b. A meta-analysis accounting for sources of variability to estimate heat resistance reference parameters of bacteria using hierarchical Bayesian modeling: Estimation of D at 121.1°C and pH 7, zT and zpH of *Geobacillus stearothermophilus*. *International Journal of Food Microbiology,* 161**,** 112-120.

ROSE, J. B. 9 November 2015. *RE: Concurrent microbial exposures.* Type to BEAUDEQUIN, D.

ROSE, J. B., HAAS, C. N., GURIAN, P. L., KOOPMAN, J. S. & CENTER FOR ADVANCING MICROBIAL RISK ASSESSMENT 2008. Instruction manual for quantitative microbial risk assessment (QMRA). East Lansing, Michigan: Michigan State University.

ROSER, D. J., PETTERSON, S. A., SIGNOR, R., ASHBOLT, N., NILSSON, P. & THORWALDSDOTTER, R. 2006. How to implement QMRA to estimate baseline and hazardous event risks with management end uses in mind. *Microrisk.* Sydney, Australia: University of New South Wales, University of Lund.

ROSS, R. 1911. *The prevention of malaria,* London, John Murray.

RYU, H. 2003. *Microbial quality and risk assessment in various water cycles in the southwestern United States* Doctor of Philosophy, Arizona State University.

SAH, L., ROUSSEAU, D. P. L. & HOOIJMANS, C. M. 2012. Numerical modelling of waste stabilization ponds: Where do we stand? *Water, Air, and Soil Pollution,* 223**,** 3155-3171.

SALES-ORTELLS, H., FERNANDEZ-CASSI, X., TIMONEDA, N., DÜRIG, W., GIRONES, R. & MEDEMA, G. 2015. Health risks derived from consumption of lettuces irrigated with tertiary effluent containing norovirus. *Food Research International,* 68**,** 70-77.

SAQQAR, M. M. & PESCOD, M. B. 1992. Modelling coliform reduction in wastewater stabilization ponds. *Water Science and Technology,* 26**,** 1667-1677.

SCHMIDT, P. J. & EMELKO, M. B. 2011. QMRA and decision-making: Are we handling measurement errors associated with pathogen concentration data correctly? *Water Research,* 45**,** 427-438.

SCHMIDT, P. J., EMELKO, M. B. & REILLY, P. M. 2010. Quantification of analytical recovery in particle and microorganism enumeration methods. *Environmental Science and Technology,* 44**,** 1705-1712.

SCHMIDT, P. J., PINTAR, K. D. M., FAZIL, A. M., FLEMMING, C. A., LANTHIER, M., LAPRADE, N., SUNOHARA, M. D., SIMHON, A., THOMAS, J. L., TOPP, E., WILKES, G. & LAPEN, D. R. 2013a. Using *Campylobacter* spp. and *Escherichia coli* data and Bayesian microbial risk assessment to examine public health risks in agricultural watersheds under tile drainage management. *Water Research,* 47**,** 3255-3272.

SCHMIDT, P. J., PINTAR, K. D. M., FAZIL, A. M. & TOPP, E. 2013b. Harnessing the theoretical foundations of the exponential and beta-Poisson dose-response models to quantify parameter uncertainty using Markov chain Monte Carlo. *Risk Analysis,* 33**,** 1677–1693.

SEIDU, R., HEISTAD, A., AMOAH, P., DRECHSEL, P., JENSSEN, P. D. & STENSTRÖM, T. A. 2008. Quantification of the health risk associated with wastewater reuse in Accra, Ghana: A contribution toward local guidelines. *Journal of Water and Health,* 6**,** 461-471.

SHIKLOMANOV, I. A. 2000. Appraisal and assessment of world water resources. *Water International,* 25**,** 11-32.

SHILTON, A. 2005. *Pond treatment technology,* United Kingdom, IWA Publishing.

SHUVA, H. & FATTAL, B. 2003. Control of pathogenic microorganisms in wastewater recycling and reuse in agriculture. *In:* MARA, D. D. & HORAN, N. J. (eds.) *Handbook of water and wastewater microbiology.* San Diego: Academic Press.

SHUVAL, H. I., LAMPERT, Y. & FATTAL, B. 1997. Development of a risk assessment approach for evaluating wastewater reuse standards for agriculture. *Water Science and Technology,* 35**,** 15-20.

SILVERMAN, A. I., NELSON, K. L., AKRONG, M. O., AMOAH, P. & DRECHSEL, P. 2013. Quantification of human norovirus GII, human adenovirus, and fecal indicator organisms in wastewater used for irrigation in Accra, Ghana. *Journal of Water and Health,* 11**,** 473-488.

SINCLAIR, M., O'TOOLE, J., GIBNEY, K. & LEDER, K. 2015. Evolution of regulatory targets for drinking water quality. *Journal of Water and Health,* 13**,** 413-426.

SINTON, L. W., HALL, C. H., LYNCH, P. A. & DAVIES-COLLEY, R. J. 2002. Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. *Applied and Environmental Microbiology,* 68**,** 1122-1131.

SMEETS, P., RIETVELD, L., HIJNEN, W., MEDEMA, G. & STENSTROM, T. A. 2006. Efficacy of water treatment processes. *Microrisk.* Delft, The Netherlands: Delft University, Kiwa Water Research, Swedish Institute for Infectious Disease Control.

SMEETS, P. W. M. H. 2013. *Stochastic modelling of drinking water treatment in quantitative microbial risk assessment,* London, IWA Publishing.

SMID, J., DE JONGE, R., HAVELAAR, A. H. & PIELAAT, A. 2013. Variability and uncertainty analysis of the cross-contamination ratios of *Salmonella* during pork cutting. *Risk Analysis,* 33**,** 1100-1115.

SMID, J. H., HERES, L., HAVELAAR, A. H. & PIELAAT, A. 2012. A biotracing model of *Salmonella* in the pork production chain. *Journal of Food Protection,* 75**,** 270-280.

SMID, J. H., SWART, A. N., HAVELAAR, A. H. & PIELAAT, A. 2011. A practical framework for the construction of a biotracing model: application to *Salmonella* in the pork slaughter chain. *Risk Analysis,* 31**,** 1434-1450.

SMID, J. H., VERLOO, D., BARKER, G. C. & HAVELAAR, A. H. 2010. Strengths and weaknesses of Monte Carlo simulation models and Bayesian belief networks in microbial risk assessment. *International Journal of Food Microbiology,* 139 (Supplement 1)**,** S57-S63.

SOLLER, J. 2008. An introduction to quantitative microbial risk assessment. *US EPA 2008 Stakeholders Meeting.* Washington DC.

SOLLER, J. A. 2012. How well does science support the QMRA process – An overview. *State-of-the-Science Symposium: Fecal Source Identification and Associated Risk Assessment Tools.* Southern California Coastal Water Research Project Authority.

SOLLER, J. A., EFTIM, S., WADE, T. J., ICHIDA, A. M., CLANCY, J. L., JOHNSON, T. B., SCHWAB, K., RAMIREZ-TORO, G., NAPPIER, S. & RAVENSCROFT, J. E. 2016. Use of quantitative microbial risk assessment to improve interpretation of a recreational water epidemiological study. *Microbial Risk Analysis,* 1**,** 2-11.

SOLLER, J. A. & EISENBERG, J. N. S. 2008. An evaluation of parsimony for microbial risk assessment models. *Environmetrics,* 19**,** 61-78.

SOLLER, J. A., OLIVIERI, A. W., CROOK, J., PARKIN, R., SPEAR, R. C., TCHOBANOGLOUS, G. & EISENBERG, J. N. S. 2003. Risk-based approach to evaluate the public health benefit of additional wastewater treatment. *Environmental Science & Technology,* 37**,** 1882-1891.

SOLLER, J. A., OLIVIERI, A. W., EISENBERG, J. N. S., SAKAJI, R. & DANIELSON, R. 2004. *Evaluation of microbial risk assessment techniques and applications,* Alexandria, Virgina, Water Environment Research Foundation.

STALEY, C., RECKHOW, K. H., LUKASIK, J. & HARWOOD, V. J. 2012. Assessment of sources of human pathogens and fecal contamination in a Florida freshwater lake. *Water Research,* 46**,** 5799-5812.

SUTER, G. W. 1999. Developing conceptual models for complex ecological risk assessments. *Human and Ecological Risk Assessment: An International Journal,* 5**,** 375-396.

SUTHERLAND, J. 1983. Normative predicates of next-generation management support systems. *IEEE Transactions on Systems, Man and Cybernetics,* SMC-13**,** 279-297.

SYMONDS, E. M., VERBYLA, M. E., LUKASIK, J. O., KAFLE, R. C., BREITBART, M. & MIHELCIC, J. R. 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 Research,* 65**,** 257-270.

TANAKA, H., ASANO, T., SCHROEDER, E. D. & TCHOBANOGLOUS, G. 1998. Estimating the safety of wastewater reclamation and reuse using enteric virus monitoring data. *Water Environment Research,* 70**,** 39-51.

TEUNIS, P., MEDEMA, G., KRUIDENIER, L. & HAVELAAR, A. 1997. Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. *Water Research,* 31**,** 1333-1346.

TEUNIS, P., MOE, C., LIU, P. & MILLER, S. 2008. Norwalk virus: How infectious is it? *Journal of Medical Virology,* 80**,** 1468-1476.

TEUNIS, P. F. M., CHAPPELL, C. L. & OKHUYSEN, P. C. 2002. *Cryptosporidium* dose-response studies: Variation between hosts. *Risk Analysis,* 22**,** 475-485.

TEUNIS, P. F. M., NAGELKERKE, N. J. D. & HAAS, C. N. 1999. Dose response models for infectious gastroenteritis. *Risk Analysis,* 19**,** 1251-1260.

THOEYE, C., VAN EYCK, K., BIXIO, D., WEEMAES, M. & DE GUELDRE, G. 2003. Methods used for health risk assessment. *In:* AERTGEERTS, R. & ANGELAKIS, A. N. (eds.) *State of the art report: Health risks in aquifer recharge using reclaimed water.* Copenhagen, Denmark: WHO Regional Office for Europe.

TOZE, S., HODGERS, L., MATTHEWS, B., STRATTON, H., AHMED, W., COLLINS, S., SCHROEDER, S. & SIDHU, J. 2012. Presence and removal of enteric microorganisms in south east Queensland wastewater treatment plants. *Technical Report No. 55.* Queensland: Urban Water Security Research Alliance.

UN-WATER. 2014. *Water quality* [Online]. Geneva, Switzerland: United Nations World Meteorological Organization. Available: http://www.unwater.org/statistics/thematic-factsheets/en/ [Accessed 19 January 2015.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY-UNITED STATES DEPARTMENT OF AGRICULTURE/FOOD SAFETY AND INSPECTION SERVICE 2012. *Microbial risk assessment guideline. Pathogenic microorganisms with focus on food and water*, Interagency Microbiological Risk Assessment Guideline Workgroup.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY 1989. National primary drinking water regulations: Filtration, disinfection, turbidity, *Giardia lamblia,* viruses, *Legionella* and heterotrophic bacteria; final rule (40 CFR Parts 141 and 142). *Federal register 54:27486.*

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY 2010. Quantitative microbial risk assessment to estimate illness in freshwater impacted by agricultural animal sources of fecal contamination. EPA 822-R-10-005. Washington, D.C.: Office of Water.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY 2014a. Microbiological risk assessment (MRA) tools, methods, and approaches for water media. EPA-820-R-14-009. Washington, DC: Office of Science and Technology Office of Water.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY 2014b. Risk assessment forum white paper: Probabilistic risk assessment methods and case studies. Washington, D.C.: Risk Assessment Forum, Office of the Science Advisor, US EPA.

UNITED STATES FOOD AND DRUG ADMINISTRATION. 2013. *Risk and safety assessment. Gaps in the knowledge base* [Online]. Available: http://www.fda.gov/Food/FoodScienceResearch/RiskSafetyAssessment/ucm262873.htm [Accessed 1 December 2014.

UNIVERSITY OF PITTSBURGH DECISION SYSTEMS LABORATORY. 2013. *GeNIe, Edition 2.0.4843.0* [Online]. Available: http://genie.sis.pitt.edu [Accessed 28 July 2015.

UUSITALO, L. 2007. Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling,* 203**,** 312-318.

UUSITALO, L., KUIKKA, S., KAUPPILA, P., SÖDERKULTALAHTI, P. & BÄCK, S. 2012. Assessing the roles of environmental factors in coastal fish production in the northern Baltic Sea: A Bayesian network application. *Integrated Environmental Assessment and Management,* 8**,** 445-455.

VAN DEN AKKER, B., WHIFFIN, V., COX, P., BEATSON, P., ASHBOLT, N. J. & ROSER, D. J. 2011. Estimating the risk from sewage treatment plant effluent in the Sydney catchment area. *Water Science and Technology,* 63**,** 1707-1715.

VARELA, A. R. & MANAIA, C. M. 2013. Human health implications of clinically relevant bacteria in wastewater habitats. *Environmental Science and Pollution Research,* 20**,** 3550-3569.

VARIS, O. & KUIKKA, S. 1999. Learning Bayesian decision analysis by doing: Lessons from environmental and natural resources management. *Ecological Modelling,* 119**,** 177-195.

VON SPERLING, M. 2007. *Waste stabilisation ponds,* London, IWA Publishing.

VORKAS, C. 1999. *Bacteriophage tracers in the identification of pathogen removal deficiencies in waste stabilisation ponds.* Doctor of Philosophy, The University of Surrey.

VOSE, D. 2000. *Risk analysis: A quantitative guide. 2nd Edition,* West Sussex, England, John Wiley & Sons.

WALTERS, C. J. 1986. *Adaptive management of renewable resources,* New York, Macmillan Publishing Co.

WALTON, M. & DEMING, W. E. 1986. *The Deming management method,* New York, U.S.A., The Berkley Publishing Group.

WANG, H., RISH, I. & MA, S. 2002. Using sensitivity analysis for selective parameter update in Bayesian network learning. *AAAI Technical Report SS-02-03*. Palo Alto, California: Association for the Advancement of Artificial Intelligence.

WATER RESEARCH AUSTRALIA. 2014a. *Log removal values in wastewater treatment. Fact sheet* [Online]. Available: www.waterra.com.au/ [Accessed 25 August 2015.

WATER RESEARCH AUSTRALIA. 2014b. Research update: Cryptosporidium toolbox. Available: http://www.waterra.com.au/publications/document-search/?download=1129 [Accessed 23 January 2015].

WEISSTEIN, E. W. 2013. *Bayesian analysis* [Online]. MathWorld--A Wolfram Web Resource. Available: http://mathworld.wolfram.com/BayesianAnalysis.html [Accessed 21/10/13.

WESTRELL, T., TEUNIS, P., VAN DEN BERG, H., LODDER, W., KETELAARS, H., STENSTRÖM, T. A. & DE RODA HUSMAN, A. M. 2006. Short- and long-term variations of norovirus concentrations in the Meuse river during a 2-year study period. *Water Research,* 40**,** 2613-2620.

WHELAN, G., KIM, K., PELTON, M. A., SOLLER, J. A., CASTLETON, K. J., MOLINA, M., PACHEPSKY, Y. & ZEPP, R. 2014. An integrated environmental modeling framework for performing quantitative microbial risk assessments. *Environmental Modelling and Software,* 55**,** 77-91.

WIDDOWSON, M.-A., SULKA, A., BULENS, S. N., BEARD, R. S., CHAVES, S. S., HAMMOND, R., SALEHI, E. D. P., SWANSON, E., TOTARO, J., WORON, R., MEAD, P. S., BRESEE, J. S., MONROE, S. S. & GLASS, R. I. 2005. Norovirus and foodborne disease, United States, 1991-2000. *Emerging Infectious Diseases,* 11**,** 95-102.

WORLD HEALTH ORGANISATION. 2006. *Guidance note for programme managers and engineers: Third edition of the guidelines for the safe use of wastewater, excreta and greywater in agriculture and aquaculture* [Online]. Available: http://www.who.int/water_sanitation_health/wastewater/human_waste/en/.

WORLD HEALTH ORGANISATION. 2011. *Enterohaemorrhagic Escherichia coli (EHEC)* [Online]. Available: http://www.who.int/mediacentre/factsheets/fs125/en/ [Accessed 31 July 2014.

WORLD HEALTH ORGANIZATION 2006. WHO guidelines for the safe use of wastewater, excreta and greywater: Volume 2. Wastewater use in agriculture. France: World Health Organization, United Nations Environment Program, Food and Agriculture Organization.

WORLD HEALTH ORGANIZATION 2008. Guidelines for drinking-water quality. Third edition incorporating 1st and 2nd addenda, Vol. 1, Recommendations. Geneva: World Health Organisation.

WU, J., LONG, S. C., DAS, D. & DORNER, S. M. 2011. Are microbial indicators and pathogens correlated? A statistical analysis of 40 years of research. *Journal of Water and Health,* 9**,** 265-278.

YANG, Z. L., BONSALL, S., WANG, J. & FANG, Q. G. 2006. Relative risk analysis using Bayesian networks and evidential reasoning. *In:* ZIO, G. S. (ed.) *Safety and Reliability for Managing Risk.* London: Taylor and Francis Group.

ZEISE, L., BOIS, F. Y., HATTIS, D., RUSYN, I. & GUYTON, K. Z. 2013. Addressing human variability in next-generation human health risk assessments of environmental chemicals. *Environmental Health Perspectives,* 121**,** 23-31.

ZWIETERING, M. H. 2009. Quantitative risk assessment: Is more complex always better? Simple is not stupid and complex is not always more correct. *International Journal of Food Microbiology,* 134**,** 57-62.

Appendices

Appendix A

Industry summary

1. Quantitative microbial risk assessment (QMRA)

The accepted methods of determining microbial safety of water and wastewater are known collectively as quantitative microbial risk assessment (QMRA). QMRA is a structured, mathematical method of assessing microbial risk and broadly comprises six steps – hazard identification, dose-response assessment, exposure assessment, risk characterisation, risk management and risk communication (Figure A1). As QMRA is a pathogen-specific method, the hazard identification step focuses on the pathogen of interest and its adverse consequences, describing the infection and disease processes in as much quantitative detail as possible. Dose-response assessment is a mathematical evaluation of the probability of an exposed individual becoming infected. Exposure assessment is based on a scenario and describes numerically, by what means, how much and how often the individual is exposed to the organism. Risk characterisation combines information from dose-response and exposure steps and describes the probability and severity of the outcome of exposure in units that can be compared with accepted benchmarks. The risk management step describes measures taken to reduce the harm caused by the hazard and can be descriptive or mathematical, depending on data availability. Risk communication, the final step of the process, conveys the results of the assessment to managers and stakeholders, and is as important as the other components of the assessment.

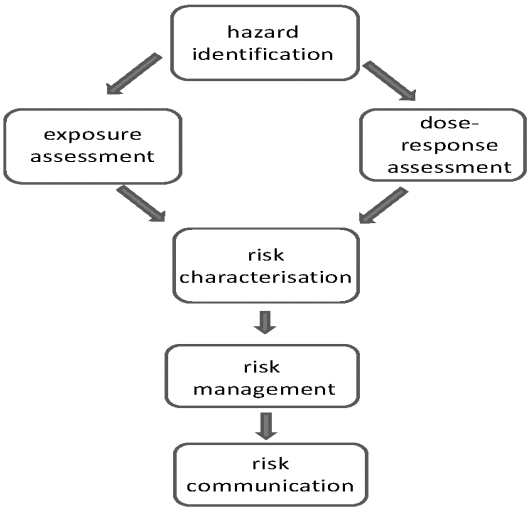

Figure A1. Generic risk assessment framework.

As QMRA is based on mathematical models, it can be confusing to many stakeholders. In addition, dependence on data that may be lacking or poor in quality, as well as variability in natural systems and uncertainty about the chosen models also contribute to its challenges. This research focuses chiefly on portraying and quantifying components of the exposure assessment and the following paragraph describes more fully the difficulties faced by risk assessors in characterising the exposure pathway for a QMRA.

The number of pathogens in an exposure is clearly an important influencing factor for infection likelihood. Therefore, in a context of water recycling, pathogen concentration in the treated wastewater is a key variable. Consider theoretically all of the variables that could affect pathogen concentration during wastewater treatment. These include treatment plant operating conditions, environmental conditions such as rainfall, wind and temperature, impacts of wildlife, and the organic and physicochemical makeup of the wastewater itself. Consider as well the difficulties in counting pathogens and indicators, due both to their size and to their low numbers in treated water. Add to that the microbial population growth and die-off that might occur during storage and distribution to the point of exposure. There are also consumption scenario factors to consider, such as how much wastewater is ingested by an individual and how often. These are the variables for which data is required to quantify a single exposure scenario. Finding accurate data for these variables to

account for every imagined exposure scenario is clearly problematic. Without well mapped, quantified exposure pathways, blanket standards are frequently used for recycled water to minimise risk, driving up treatment costs and inhibiting uptake of reuse schemes.

2. Bayesian networks (BNs)

In this study, Bayesian networks (BNs) are used as a complementary approach to overcome some of the challenges presented by QMRA. BNs are graphical models, based on probability theory. Figure A2 is an example of a simple BN. Arrows connect the nodes representing variables such as temperature, with their direction indicating causal influence of other variables. For example, the network in Figure A2 indicates that microbial growth is influenced by temperature and competitor population. During design of the network, each node (variable) is assigned a number of categories referred to as 'states'. In the BN in Figure A2 for example, the node *Temperature* is assigned the states high and low. A probability is assigned to each state, indicating the likelihood that the variable might be in that state (e.g., high or low). These probabilities can come from the literature, fieldwork, simulation models, or expert opinion. If no information exists, equal probabilities can be assigned to the states, e.g., 0.5 each for high and low.

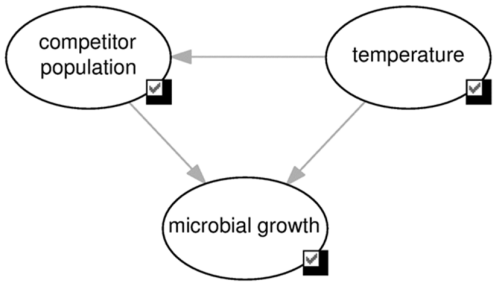

Figure A2. Simple Bayesian network of factors influencing microbial growth.

In Figure A2, the *Temperature* node is a 'root' node, meaning it has no other variables influencing it. The probability table underlying the *Temperature* node might look like Table A1:

Table A1

Potential probabilities for Temperature node - Bayesian network for microbial growth

temperature	
high	**low**
0.2	0.8

The *Competitor population* node is influenced by temperature, so the probabilities underlying this node are referred to as 'conditional' probabilities. To illustrate, the probability of *Competitor population* being high depends upon whether *Temperature* is high or low, because according to the model, *Temperature* influences *Competitor population*. For this particular competitor population, high temperatures are favourable for growth, so when *Temperature* is high, it is almost a certainty (P = 0.9) that *Competitor population* will be high. The table of conditional probabilities underlying the *Competitor population* node might look like Table A2.

Table A2

Potential conditional probability table for Competitor population node – Bayesian network for microbial growth

		temperature	
		high	**low**
competitor	high	0.9	0.1
population	low	0.1	0.9

From the model in Figure A2, it can be seen that *Microbial growth* is influenced by both *Temperature* and *Competitor population*. Table A3 displays the probabilities for *Microbial growth*, conditional upon the states in the nodes *Temperature* and *Competitor population*.

Table A3

Potential conditional probability table for Microbial growth node – Bayesian network for microbial growth

		microbial growth	
temperature	**competitor population**	**low**	**high**
low	low	0.65	0.35
low	high	0.99	0.01
high	low	0.01	0.99
high	high	0.45	0.55

From Table A3, it can be seen that when temperature is high and competitor population is low, the probability of microbial growth being high is 0.99 - almost a certainty.

3. How conditional probabilities for BNs were derived from QMRA data in this study

The conditional probabilities for the BNs in this study have been calculated from data generated by QMRA models. The input data for the QMRA models (such as pathogen concentration) and the threshold values for node states (e.g., threshold for 'low' pathogen concentration) were found in the literature. The following steps describe the way in which QMRA was used to populate the nodes of the BNs. This is described in detail in Chapter 6. For the purposes of illustrating how conditional probabilities for the BN nodes were calculated from QMRA data, Figure A3 represents the factors influencing *Cryptosporidium* oocyst concentration in primary treated wastewater.

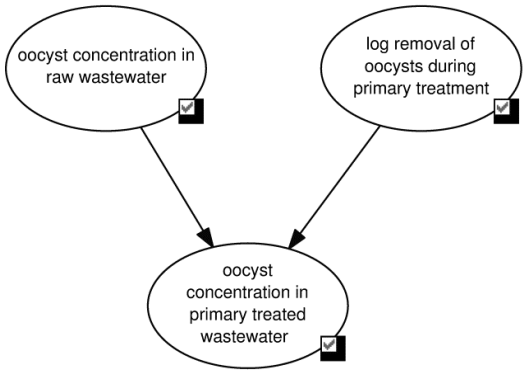

Figure A3. Simple Bayesian network representing the factors influencing *Cryptosporidium* oocyst concentration in primary treated wastewater.

1. Data points from the QMRA models were labelled 'high' or 'low' using chosen threshold values, as illustrated in Table A4:

Table A4

Data points from QMRA models transformed to Bayesian network node 'states' based on chosen threshold values

oocyst concentration in raw wastewater (#/mL)	**log removal in primary treatment**	**oocyst concentration in primary treated wastewater (#/mL)**	**oocyst concentration in raw wastewater (#/mL)**	**log removal in primary treatment**	**oocyst concentration in primary treated wastewater (#/mL)**
low: ≤ 3.68; high: >3.68	**low: ≤ 0.25; high: > 0.25**	**low: ≤ 2.04; high: >2.04**			
7.17	0.41	2.76	H	H	H
2.97	0.45	1.06	L	H	L
3.15	0.30	1.59	L	H	L
6.41	0.01	6.31	H	L	H
4.28	0.27	2.28	H	H	H

2. Root nodes in the BN (e.g., *Oocyst concentration in raw wastewater* and *Log removal of oocysts during primary treatment* in Figure A3) were allocated equal probabilities for all states, i.e., 0.5 for high and 0.5 for low.

3. Conditional probabilities for child nodes in the BN (e.g., the node *Oocyst concentration in primary treated wastewater*) were calculated in Excel using pivot tables and a formula known as Bayes' rule. Pivot tables for selected variables summarise the number of times state combinations occur (e.g., for the variables shown in Table A4, high-high-high, low-high-low, etc.). Bayes' rule is then used to calculate the conditional probability values from the pivot table counts. The result is a conditional probability table for every child node in the network, as shown in the example in Table A5. Once these probabilities are entered into a network for each node, it is ready to use.

Table A5

Conditional probability table underlying the node Oocyst concentration during post primary treatment node, which is influenced by the nodes Oocyst concentration in raw wastewater and the Log removal of oocysts during primary treatment

Log removal of oocysts during primary treatment		**low**		**high**	
Oocyst concentration in raw wastewater		**low**	**high**	**low**	**high**
Oocyst concentration in primary treated wastewater	low	0.6751	0	1	0.3242
	high	0.3249	1	0	0.6758

4. How do BN nodes interact?

After the probabilities for the node states are entered into the network, the user 'compiles' the network, usually with a single mouse click. This instructs the software to run its underlying mathematical processes, establishing the baseline status of the network. Following this, 'new evidence' can be entered into one or more nodes to model various scenarios. For example, changing the node *Oocyst concentration in raw wastewater* to 100% 'high' from its baseline status (again with a mouse click) can simulate an event that would result in such an increase. After the introduction of new evidence, the user re-compiles the network and the software recalculates the likelihoods of the node states, in the light of the new evidence. The effect of the new evidence is propagated through the network (backwards as well as forwards) and all of the nodes in the network will then change to reflect how the updated information has affected them. This ability of BN models to propagate information in both directions is one of their most useful features.

5. What do BNs add to QMRA?

BNs can be used with QMRA in a number of ways. This thesis, in which QMRA-based BNs are used to illustrate, quantify and manipulate components of exposure pathways, is one example of how BNs and QMRA can be used together. This research has found that BNs effectively augment the QMRA methodology, making it more transparent, flexible and accessible to stakeholders and decision makers. The unique features of BNs that complement the QMRA method are:

- **BNs are able to support 'backwards' inference**; this enables discovery of the key drivers for an outcome of interest. Alternatively, the desired outcome can be set (introduced as evidence) in a target node, to determine conditions required 'upstream' to achieve the desired outcome.
- **BNs enable multiple interacting variables to be changed at one time.** In scenario or 'what if' analyses in a BN, variable states are 'changed' with a single mouse click, simulating changed conditions in the system being modelled. The outcome nodes in a BN respond immediately to changes in 'upstream' variables.
- **BNs are visual models**; multiple stakeholders with differing levels of knowledge can look at BN models and understand their responses. BNs also

promote engagement; whereas the presentation of QMRA's mathematical models might 'turn off' many stakeholders, the visual appeal of a BN keeps them interested and engaged.

- **BNs offer a transparent evidence base** to inform management options. Managers can see at a glance which treatment option or risk reduction measure will have the most impact, and the probability distributions for the variables.
- **BNs can propagate 'knowledge' throughout the network** in both directions, enabling updating of poor quality prior information in response to new evidence. This is particularly useful in QMRA, where data may be poor in quality due to the inherent difficulties of microorganism enumeration.
- **BNs have an instant updating capability**, enabling rapid appraisal of scenarios involving simultaneous changes in multiple variables. The impact of a change in one variable is spread through all variables in the network instantaneously by underlying software processes and resulting changes in nodes are immediately visible.
- **BNs represent uncertainty clearly** at variable level via visual probability distributions.
- **BNs transform risk benchmarks** from a dichotomous scheme (e.g., 'tolerable/not tolerable') to a probability continuum (e.g., '94% chance of achieving tolerable risk of infection').

6. Examples of how the BNs in this research can be used to support decision making in a water recycling context.

The outcomes of this study enable practitioners to selectively determine treatment types and risk management options to optimise treatment and reduce risk. The following potential applications of the models developed in the study demonstrate their utility for practitioners such as water utilities managers, regulatory authorities, treatment plant operators, risk modellers and public facility managers:

- A sewage treatment plant operator might use sensitivity analysis in a model to determine which step in the treatment chain has the most impact on pathogen concentration;

- Hazardous event conditions such as power or pump failure could be simulated in the model by a plant operator or manager to determine likely risk outcomes if water is reused in a particular scenario;
- A regulator might enact a guideline stating that under normal circumstances an 80% chance of achieving a tolerable disease burden is acceptable for a pathogen in a particular irrigation scenario, but under specified extreme conditions a 70% chance of achieving a tolerable disease burden would be accepted;
- A regulator might use a model to justify a lower class of water being used to irrigate edible crops, if a 3 day irrigation period was applied or crop washing was undertaken;
- A water utilities manager could model particular water treatment conditions in conjunction with various exposure profiles to determine the type of public open space for which the water is most suited, for irrigation purposes;
- A desired risk outcome, along with other variable constraints, could be set by a water utilities manager working with a plant operator to determine the conditions required in the treatment chain to achieve the designated outcome;
- A manager of a public facility (e.g., municipal park) might use sensitivity analysis in the model to determine which onsite risk reduction measure is most efficient;
- A water scientist could use the model to indicate which variables in the model are optimal for further research expenditure in terms of improving accuracy of risk estimates.

Appendix B

Table B1

Summary of applications of Bayesian networks in quantitative microbial risk assessment

reference	domain	knowledge source informing model structure	source of conditional probability table values	number of nodes	software	model validation method	belief updating	quantifies hazard	prediction	separation of uncertainty and variability	uncertainty reduction reported	scenario assessment and decision making	software developed	new method
Barker et al. (2002)	QRA for food-borne pathogen - hazards arising from *Clostridium botulinum* growth and toxin production	literature, expert knowledge	empirical data, model simulations	10 (spore concentrations), 23 (bacterial growth)	Hugin	sensitivity analysis	data	y	y	not discussed	y	y	n	n
Pouillot et al. (2003)	QRA for food-borne pathogen - *Listeria monocytogenes* in milk	literature, expert knowledge	model simulations	29	WinBUGS	based on existing models	data, maximum likelihood estimates	y	y	y	y	not discussed	n	y
Parsons et al. (2005)	QRA for a food-borne pathogen - *Salmonella* spp. in poultry meat production chain	expert knowledge	literature, expert opinion, other unpublished sources	23 nodes in Figure 1, but 'the final model consisted of 60 parameters'.	1. NETICA 2. WinBUGS	output compared with survey data, sensitivity analysis	1. data 2. data	y	y	1. n 2. y	1. y 2. y	y	n	n

reference	domain	knowledge source informing model structure	source of conditional probability table values	number of nodes	software	model validation method	belief updating	quantifies hazard	prediction	separation of uncertainty and variability	uncertainty reduction reported	scenario assessment and decision making	software developed	new method
Delignette-Muller et al. (2006)	QRA for a food-borne pathogen - exposure assessment for Listeria monocytogenes in cold-smoked salmon	literature, observations, expert knowledge	not discussed	L. monocytogenes model: 19 nodes including 7 hyperparameters and 2 covariates; food flora model: 18 nodes including 9 hyperparameters and 2 covariates	WinBUGS	data	not discussed	y	y	y	y	not discussed	n	n
Albert et al. (2008)	QRA for food-borne pathogens – estimating the probability of campylobacterosis caused by home consumption of chicken meat	expert knowledge	model simulations, expert opinion	24 (augmented core model)	JAGS and Winbugs	data, expert evaluation, sensitivity analysis	data	y	y	not discussed	not discussed	y	n	y

reference	domain	knowledge source informing model structure	source of conditional probability table values	number of nodes	software	model validation method	belief updating	quantifies hazard	prediction	separation of uncertainty and variability	uncertainty reduction reported	scenario assessment and decision making	software developed	new method
Smid et al. (2012); Smid et al. (2011)	QRA for food-borne pathogen - Salmonella in the pork slaughter chain	literature, observations, expert knowledge	empirical data, expert opinion, literature, model simulations	63	Hugin	sensitivity analysis, based on existing models	data	y	y	not discussed	y	y	y	n
Rigaux et al. (2012a)	QRA for food-borne pathogen - Bacillus cereus in courgette puree	not discussed	not discussed	58	JAGS 2.1.0.	data	MCMC algorithm	y	y	y	y	y	n	y
Rigaux et al. (2012b)	QRA for food-borne pathogen - Geobacillus stearothermophilus in the spoilage of canned foods	literature (peer reviewed and unpublished data)	not discussed	basic model: 8; intermediate model: 10; complete model: 14	Jags 3.2.0	10-fold cross-validation	model simulations using MCMC	y	y	y	n	not discussed	n	n
Smid et al. (2013)	QRA for food-borne pathogen - Salmonella in pork production chain	not discussed	empirical data	21 variables	Hugin	data	data - sequential adaptation	y	y	n	y	not discussed	n	n
Barker and Gomez-Tome (2013)	QRA for food-borne pathogens – enterotoxigenic Staphylococcus aureus in milk	not discussed	empirical data, (published and observed), expert opinion	35 parameters	HUGIN	checked model output with some published data	not discussed	y	y	y	not discussed	y	n	n

reference	domain	knowledge source informing model structure	source of conditional probability table values	number of nodes	software	model validation method	belief updating	quantifies hazard	prediction	separation of uncertainty and variability	uncertainty reduction reported	scenario assessment and decision making	software developed	new method
Donald et al. (2009)	estimating potential health risks associated with recycled water	expert knowledge	expert opinion	14	Netica and Hugin (model 1); WinBUGS (model 2)	sensitivity analysis	expert opinion, data	y	y	n	n	y	n	y
Gronewold et al. (2011)	assessment of the potential threat of faecal contamination in surface water	literature, observations, model simulations	not discussed	17	WinBUGS	compared conventional regression analysis, 'leave one out; cross-confirmation procedure	data	y	y	y	y	y	n	y
Goulding et al. (2012)	environmental engineering/public health - assessment of public health risk from wet weather sewer overflows	literature	empirical data (published and observed), expert opinion, modelling	14	not discussed	expert evaluation, sensitivity analysis	data	y	y	n	n	y	n	n
Staley et al. (2012)	QRA for waterborne pathogens in a freshwater lake	expert knowledge, machine learning	empirical data	6	Hugin	not discussed	data	y	y	n	not discussed	y	n	n

Appendix C

Table C1

Model input parameters and distributions

variable	units	distribution[a] or point estimates	references
wastewater retained on lettuce	mL/g	normal (0.108, 0.019), truncate at 0	(Barker, 2014a, Hamilton et al., 2006, Shuval et al., 1997)
lettuce consumed	g/person/day	triangular (25, 40, 100)	(NRMMC-EPHC-AHMC, 2006)
pathogen die-off rate [b]	per day	normal (1.07, 0.07)	(Barker, 2014a, Petterson et al., 2001, Petterson et al., 2002)
irrigation withholding period	days	uniform (0, 1, 3)	
pathogen reduction (washing)	$\log_{10}$ units	lognormal (0.694, 0.459) truncated at 0.1 and 2.25	(Barker, 2014a)
norovirus concentration in untreated wastewater	PCR units/mL	triangular (1×10^4, 5×10^6, 1×10^7)	(Barker, 2014a)
viral $\log_{10}$ reduction during treatment	$\log_{10}$	6	(NRMMC-EPHC-AHMC, 2006)
norovirus concentration in treated wastewater	PCR units/mL	lognormal (5.005, 5.005)	as described in section 5.2.1
illness to infection ratio		uniform (0.8, 1.0)	(Atmar, 2010, Barker et al., 2013b)
annual frequency of lettuce consumption	# times/year	triangular (10, 70, 500)	(NRMMC-EPHC-AHMC, 2006)

[a]distributions were defined for root nodes in underlying model: normal (mean, standard deviation); triangular (minimum, most likely, maximum); lognormal (mean, standard deviation) uniform (minimum, median, maximum)

[b]it is assumed that the rate of die-off is constant over time, however the rate of die-off may be biphasic or multiphasic (O'Toole, 2011)

Table C2

Variable states and ranges

variable	units	references[a]	states and ranges
wastewater retained on lettuce	mL/g	(Shuval et al., 1997)	low: <0.09, medium: 0.09-0.1250, high: >0.1250
lettuce consumption	g/person/day	(NRMMC-EPHC-AHMC, 2006)	low: <40, medium: 40-70, high: >70
pathogen die-off rate	per day	(O'Toole, 2011, Petterson et al., 2001, Petterson et al., 2002)	low:<1, medium: 1-1.14, high:>1.14
irrigation withholding period	days		low: 0, medium: 1, high: 3
norovirus concentration in treated wastewater	PCR units/mL	(Atmar et al., 2014)	low:<3, medium: 3-7, high: >7
wastewater ingested	mL		low:<3, medium: 3-8, high:>8
pathogen reduction (washing)	latent variable	(Barker et al., 2014)	low: <0.03, medium: 0.03-0.08, high: >0.08
pathogen reduction (withholding irrigation)	latent variable		low: <0.2, medium: 0.2-0.8, high: >0.8
unmitigated norovirus dose	PCR units	(Teunis et al., 2008)	low:<10, medium:10-30, high: >30
total pathogen reduction	latent variable		low:<0.005, medium: 0.005-0.0325, high:>0.0325
mitigated norovirus dose	PCR units	(Teunis et al., 2008)	low:<1, medium:1-5, high:>5
risk of norovirus infection	per person per day		low:< ≤ 0.002, medium:0.002 - 0.006, high:>0.006
illness to infection ratio		(Atmar, 2010, Barker et al., 2013b)	low:<0.8667, medium: ≥0.8667, <0.9333, high: ≥0.9333
annual frequency of lettuce consumption	# times/year	(NRMMC-EPHC-AHMC, 2006)	low: ≥12, <120 times per year, medium: ≥120, <270 times per year, high: ≥270, <365 times per year, very high: ≥ 365 times per year, capped at 500
annual risk of norovirus infection	per person per year	(Mara and Sleigh, 2010)	tolerable: ≤ 0.0014, low: >0.0014, ≤0.2510, medium: >0.2510, ≤0.5007, high: >0.5007, ≤0.75035, very high:>0.75035
annual risk of illness related to norovirus	per person per year	(Mara and Sleigh, 2010)	tolerable: ≤0.0011, low: >0.0011, ≤0.2508, medium: >0.2508, ≤0.50055, high: >0.5005, ≤0.750275, very high: >0.750275

[a]thresholds and ranges for states were selected on the basis of indicative or typical values for Australian conditions

Appendix D

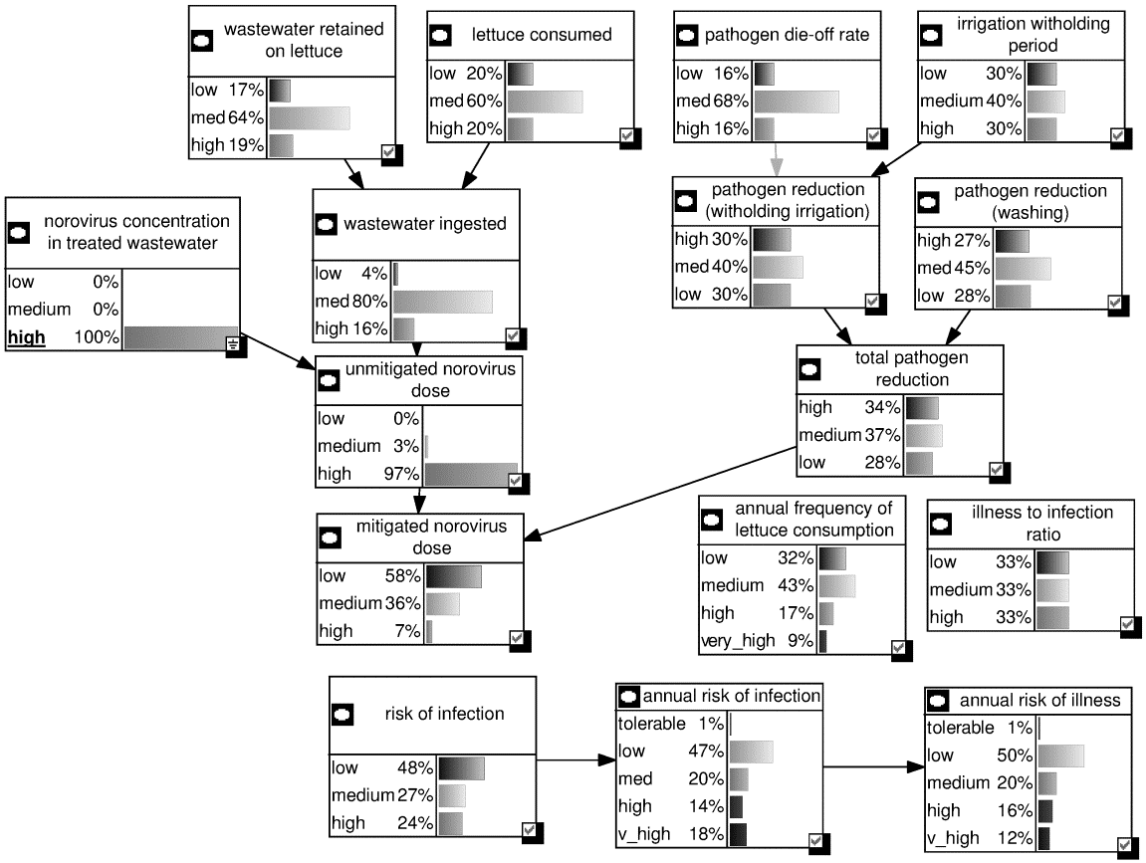

Figure D1. Bayesian network for risk of norovirus infection - scenario 'Outbreak'.

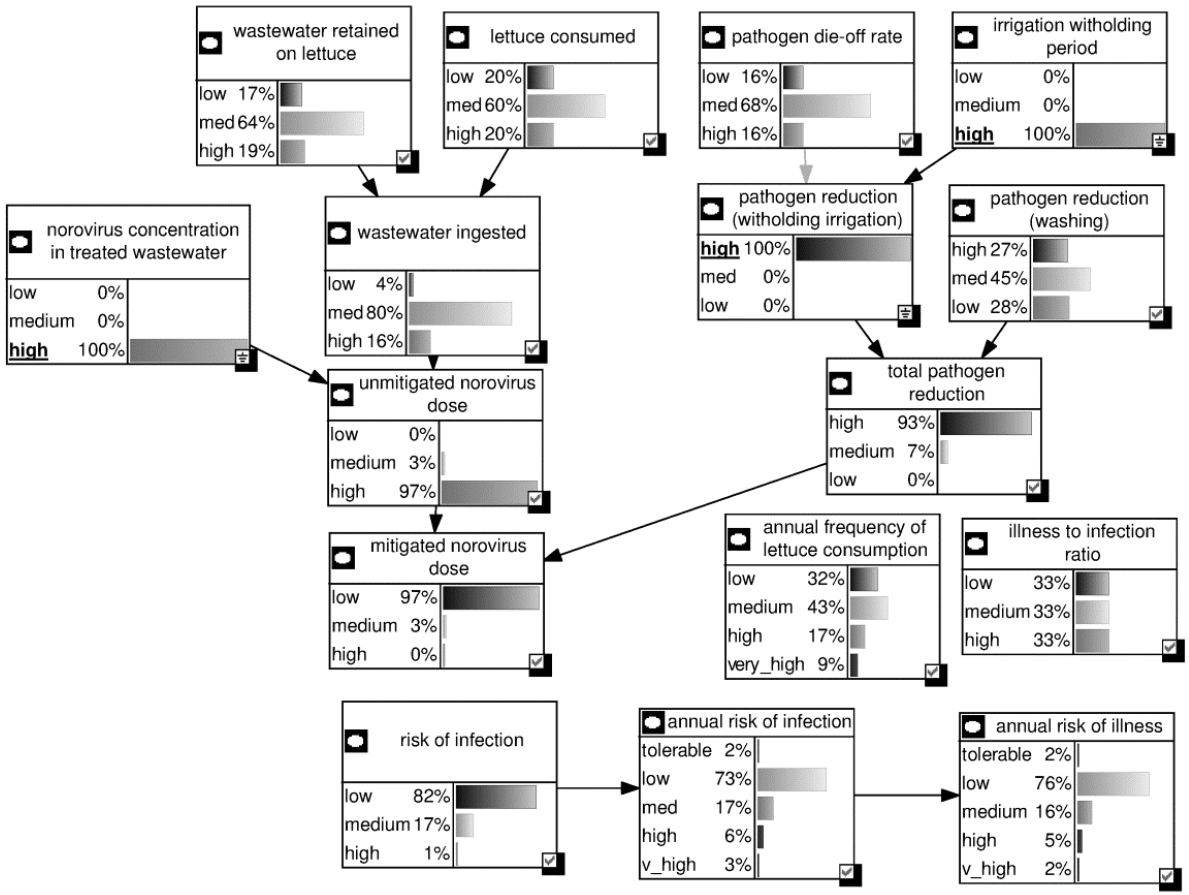

Figure D2. Bayesian network for risk of norovirus infection - scenario 'Outbreak with risk mitigation'.

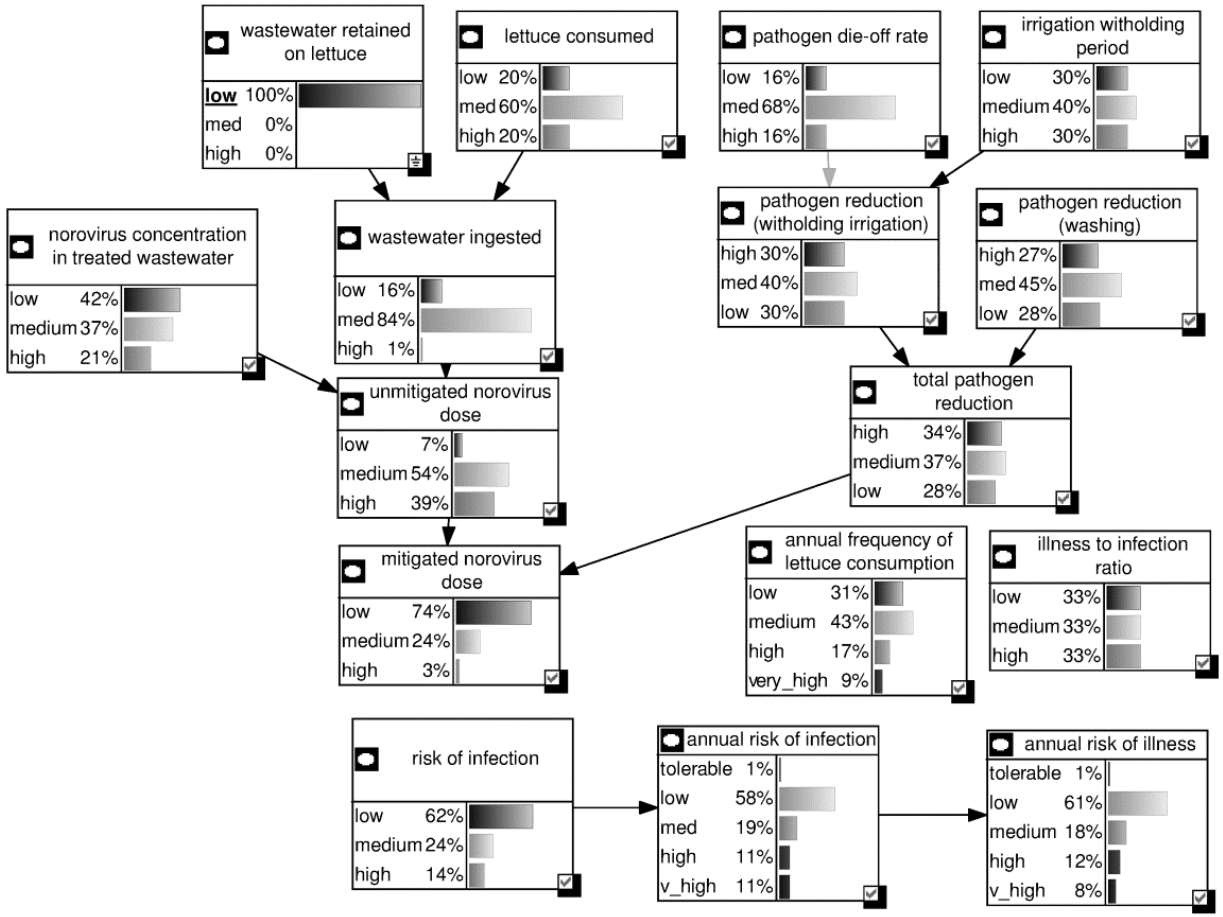

Figure D3. Bayesian network for risk of norovirus infection - scenario 'Furrow system'.

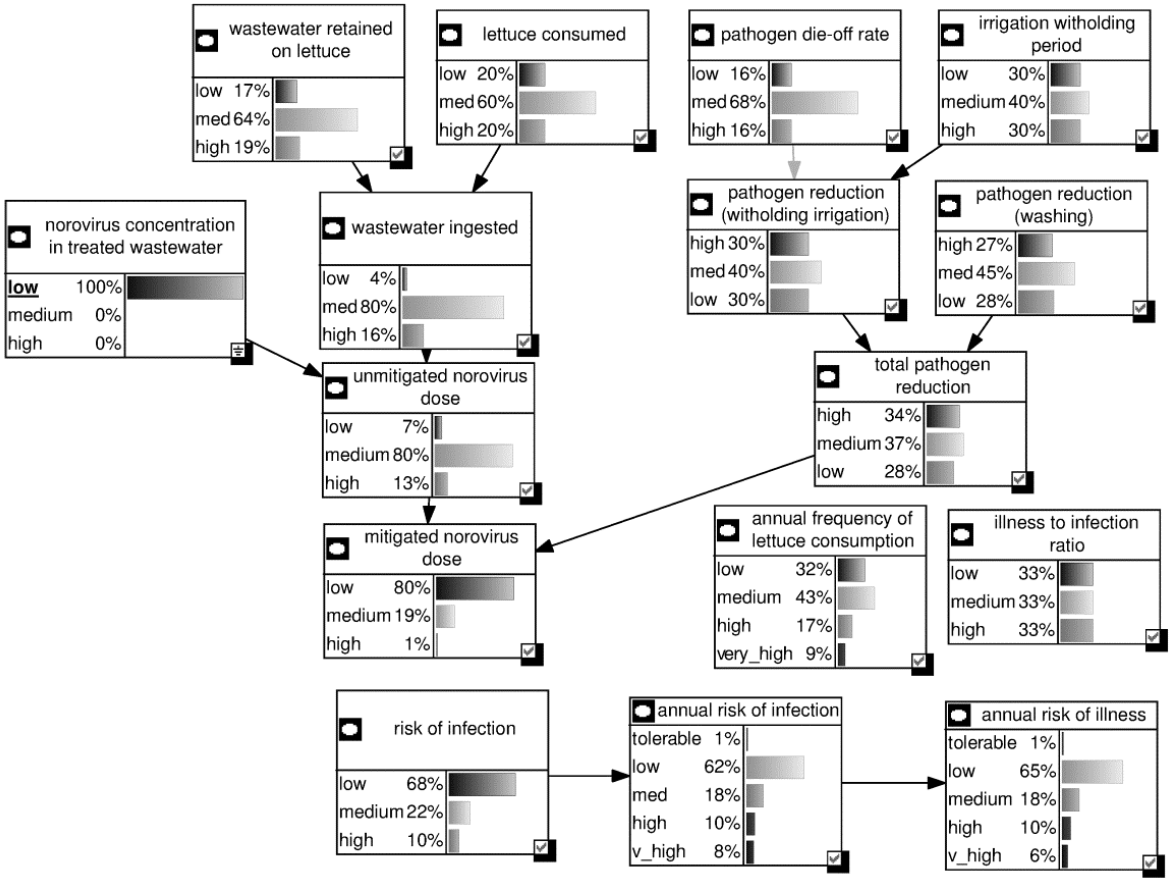

Figure D4. Bayesian network for risk of norovirus infection - scenario 'Treatment change'.

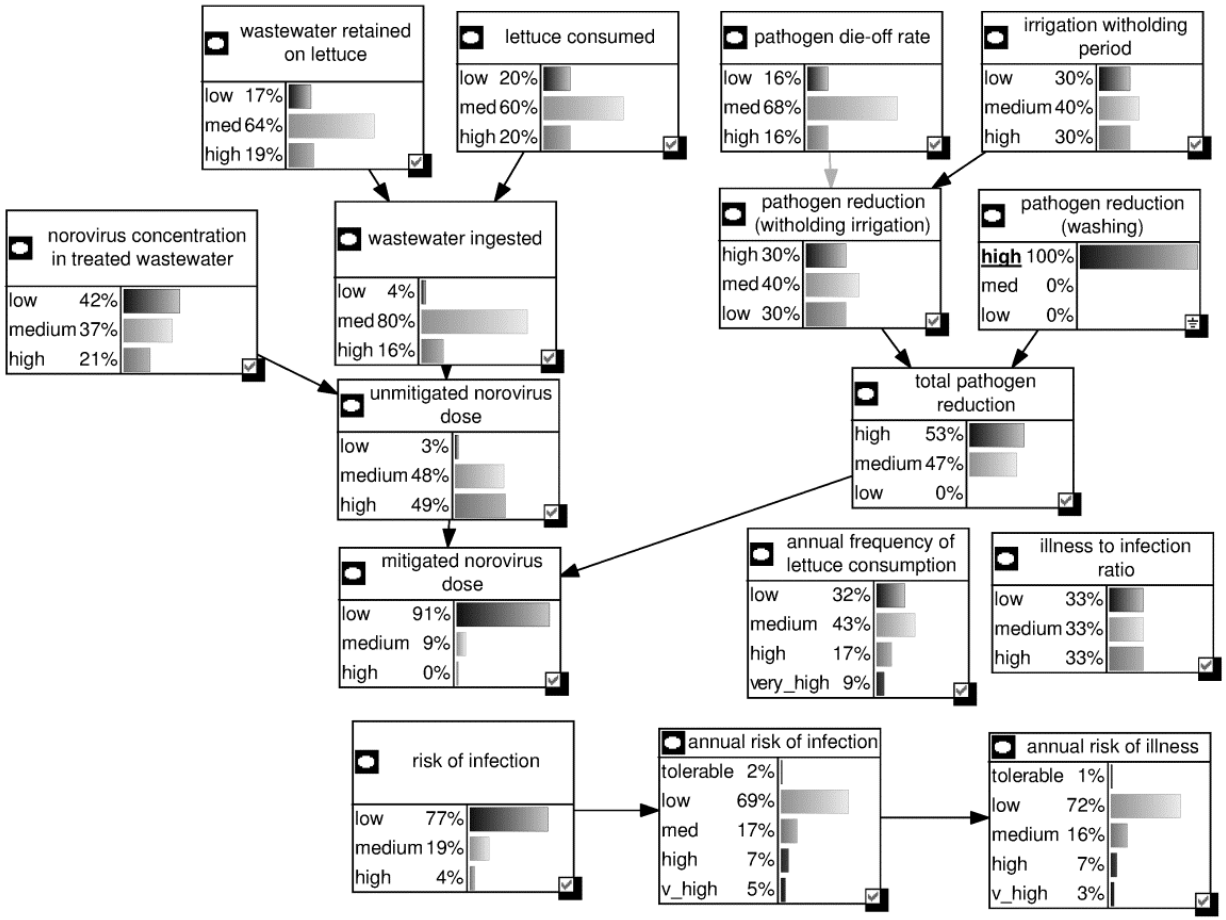

Figure D5. Bayesian network for risk of norovirus infection - scenario 'Lettuce washing'.

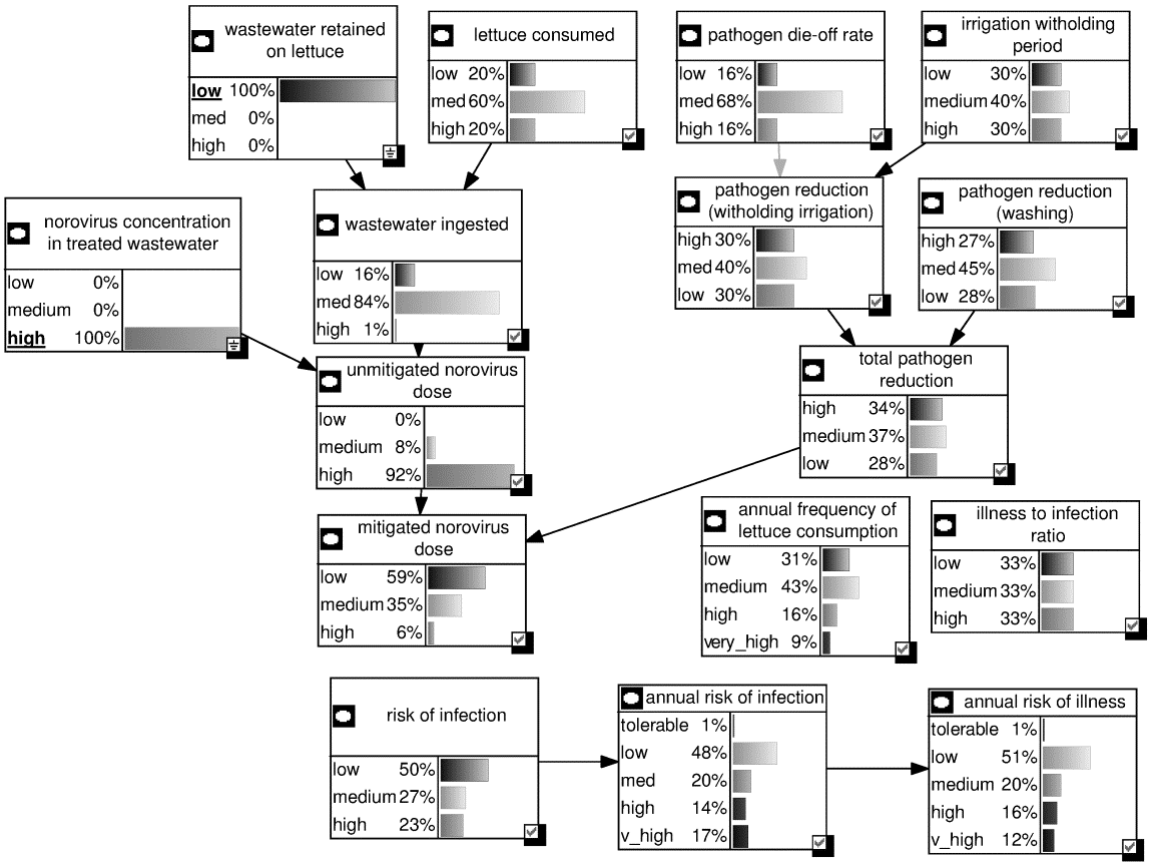

Figure D6. Bayesian network for risk of norovirus infection - scenario 'Rain'.

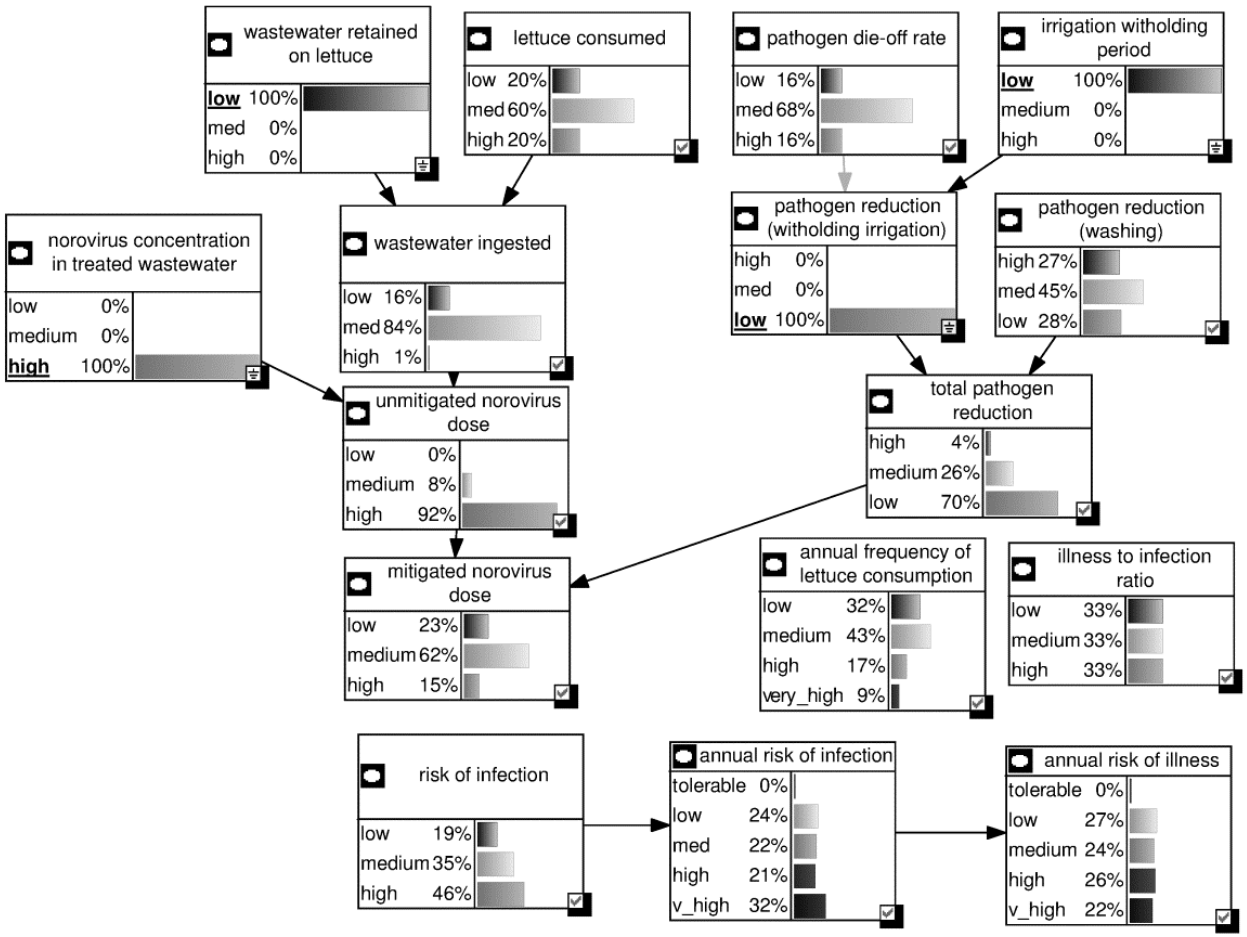

Figure D7. Bayesian network for risk of norovirus infection - scenario 'Rain with decreased withholding period'.

Appendix E

Table E1

QMRA process model input parameters and distributions - Cryptosporidium

variable	units	distribution[a] or point estimates [mean]	references
oocyst concentration in raw wastewater	oocysts/mL	triangular (0, 2, 10)	(Bartrand et al., 2013, Cunliffe, 2006, NRMMC-EPHC-AHMC, 2006, Van Den Akker et al., 2011)
log removal of oocysts during primary treatment	log_{10} units	uniform (0, 0.5)	(NRMMC-EPHC-AHMC, 2006)
log removal of oocysts during secondary treatment	log_{10} units	uniform (0.5, 1.0)	(NRMMC-EPHC-AHMC, 2006)
log removal of oocysts during lagoon storage	log_{10} units	uniform (1.0, 3.5)	(NRMMC-EPHC-AHMC, 2006)
log removal of oocysts during wetlands surface flow	log_{10} units	uniform (0.5, 1.0)	(NRMMC-EPHC-AHMC, 2006)
log removal of oocysts during wetlands subsurface flow	log_{10} units	uniform (0.5, 1.0)	(NRMMC-EPHC-AHMC, 2006)
log removal of oocysts during chlorination	log_{10} units	discrete uniform (0, 0.5)	(NRMMC-EPHC-AHMC, 2006)
log removal of oocysts due to spray drift control	log_{10} units	discrete uniform (0, 1)	(NRMMC-EPHC-AHMC, 2006)
log removal of oocysts due to 4 hr withholding of irrigation	log_{10} units	0	(Hutchison et al., 2005, Jenkins et al., 2013)
wastewater volume ingested	mL	discrete uniform (1, 5)	(Asano et al., 1992, NRMMC-EPHC-AHMC, 2006, Ryu, 2003)
frequency of visits	visits per year	discrete uniform (26, 240)	
illness to infection ratio		0.7	(Havelaar and Melse, 2003, NRMMC-EPHC-AHMC, 2006)
disease burden	DALYs per case of illness	1.5×10^{-3}	(Havelaar and Melse, 2003, NRMMC-EPHC-AHMC, 2006)
susceptibility fraction		1	(Havelaar and Melse, 2003, NRMMC-EPHC-AHMC, 2006)

[a]distributions were defined as: triangular (minimum, most likely, maximum); uniform (minimum, maximum); discrete uniform (minimum, maximum)

Table E2

Variable states and ranges - Cryptosporidium

variable	units	states and ranges	discretisation, references[a]
oocyst concentration in raw wastewater	#/mL	low: ≤ 3.68; high: >3.68	equal probabilities
log removal of oocysts during primary treatment	log_{10} units	low: ≤ 0.25; high: >0.25	equal probabilities
oocyst concentration in primary treated wastewater	#/mL	low: ≤ 2.04; high: >2.04	equal probabilities
log removal of oocysts during secondary treatment	log_{10} units	low: ≤ 0.75; high: >0.75	equal probabilities
oocyst concentration in secondary treated wastewater	#/mL	low: ≤ 0.36; high: >0.36	equal probabilities
log removal of oocysts during lagoon storage	log_{10} units	low: ≤ 2.25; high: >2.25	equal probabilities
post lagoon oocyst concentration	#/mL	low: ≤ 1.91×10^{-3}; high: >1.91×10^{-3}	equal probabilities
log removal of oocysts during wetlands surface flow	log_{10} units	low: ≤ 0.75; high: >0.75	equal probabilities
oocyst concentration post wetlands surface flow	#/mL	low: ≤ 3.37×10^{-4}; high: >3.37×10^{-4}	equal probabilities
log removal of oocysts during wetlands subsurface flow	log_{10} units	low: ≤ 0.75; high: > 0.75	equal probabilities
oocyst concentration post wetlands subsurface flow	#/mL	low: ≤ 5.95×10^{-5}; high: > 5.95×10^{-5}	equal probabilities
log removal of oocysts during chlorination	log_{10} units	on: 0.25; off: 0	equal probabilities
oocyst concentration post chlorination	#/mL	low: ≤ 4.51×10^{-5}; high: > 4.51×10^{-5}	equal probabilities
log removal of oocysts due to spray drift control	log_{10} units	on: 1; off: 0	(NRMMC-EPHC-AHMC, 2006)
log removal of oocysts due to 4 hour withholding of irrigation	log_{10} units	0	(Hutchison et al., 2005, Jenkins et al., 2013)
onsite oocyst concentration	#/mL	low: ≤ 1.4×10^{-5}; high: > 1.4×10^{-5}	equal probabilities

variable	units	states and ranges	discretisation, references[a]
wastewater volume ingested	mL	one: 1; five: 5	(Asano et al., 1992, NRMMC-EPHC-AHMC, 2006, Ryu, 2003)
dose	# oocysts	low: $\leq 3.2 \times 10^{-5}$; high: $> 3.2 \times 10^{-5}$	equal probabilities
risk of infection		low: $\leq 1.84 \times 10^{-6}$; high: $> 1.84 \times 10^{-6}$	equal probabilities
frequency of visits	visits per year[b]	fortnightly: 26; weekly: 52; twice weekly: 104; daily: 240	
annual risk of infection	pppy	tolerable: $\leq 2.2 \times 10^{-3}$; high $> 2.2 \times 10^{-3}$	(World Health Organisation, 2006)
annual risk of illness	pppy	tolerable: $\leq 6.7 \times 10^{-4}$; high $> 6.7 \times 10^{-4}$	(World Health Organisation, 2006)
annual disease burden	DALYs pppy	tolerable: $\leq 1.0 \times 10^{-6}$; high $> 1.0 \times 10^{-6}$	(World Health Organization, 2006)

[a]where indicated, thresholds and ranges for states were derived from published values

[b]for occupational exposures, 'daily' exposure is assumed to be 5 days a week for 48 weeks a year

Table E3

QMRA process model input parameters and distributions - norovirus

variable	units	distribution[a] or point estimates [mean]	references
norovirus concentration in raw wastewater	PCR units/mL	triangular (1.0×10^4, 5.0×10^6, 1.0×10^7)	(Barker, 2014a)
log removal of norovirus during primary treatment	$\log_{10}$ units	uniform (0, 0.1)	(NRMMC-EPHC-AHMC, 2006)
log removal of norovirus during secondary treatment	$\log_{10}$ units	uniform (0.5, 2)	(NRMMC-EPHC-AHMC, 2006)
log removal of norovirus during lagoon storage	$\log_{10}$ units	uniform (1, 4)	(NRMMC-EPHC-AHMC, 2006)
log removal of norovirus during wetlands surface flow	$\log_{10}$ units	uniform (1.94, 1.98)	(Gerba et al., 2013)
log removal of norovirus during wetlands subsurface flow	$\log_{10}$ units	uniform (1.94, 1.99)	(Gerba et al., 2013)
log removal of norovirus during chlorination	$\log_{10}$ units	discrete uniform (0, 2)	(NRMMC-EPHC-AHMC, 2006)
log removal of norovirus due to spray drift control	$\log_{10}$ units	discrete uniform (0, 1)	(NRMMC-EPHC-AHMC, 2006)
log removal of norovirus due to 4 hour withholding of irrigation	$\log_{10}$ units	discrete uniform (0, 0.4)	(Page et al., 2014)
wastewater volume ingested	mL	discrete uniform (1, 5)	(Asano et al., 1992, NRMMC-EPHC-AHMC, 2006, Ryu, 2003)
frequency of visits	visits per year	discrete uniform (26, 240)	
illness to infection ratio		0.67	(Atmar, 2010)
disease burden	DALYs per case of illness	uniform (3.71×10^{-4}, 6.23 × 10−3) [3.30×10^{-3}]	(Havelaar and Melse, 2003)
susceptibility fraction		uniform (0.8, 1.0) [0.9]	(Atmar, 2010, Barker et al., 2013b, Barker et al., 2013a)

[a]distributions were defined as: triangular (minimum, most likely, maximum); uniform (minimum, maximum); discrete uniform (minimum, maximum)

Table E4

Variable states and ranges - norovirus

variable	units	states and ranges	discretisation, references[a]
norovirus concentration in raw wastewater	PCR units/mL	low: ≤ 5.0 x 10^6; high: >5.0 x 10^6	equal probabilities
log removal of norovirus during primary treatment	log_{10} units	low: ≤ 0.05; high: >0.05	equal probabilities
norovirus concentration in primary treated wastewater	PCR units/mL	low: ≤ 4.44 x 10^6; high: >4.44 x 10^6	equal probabilities
log removal of norovirus during secondary treatment	log_{10} units	low: ≤ 1.25; high: >1.25	equal probabilities
norovirus concentration in secondary treated wastewater	PCR units/mL	low: ≤ 2.23 x 10^5; high: >2.23 x 10^5	equal probabilities
log removal of norovirus during lagoon storage	log_{10} units	low: ≤ 2.5; high: >2.5	equal probabilities
post lagoon norovirus concentration	PCR units/mL	low: ≤ 739; high: >739	equal probabilities
log removal of norovirus during wetlands surface flow	log_{10} units	low: ≤ 1.96; high: >1.96	equal probabilities
norovirus concentration post wetlands surface flow	PCR units/mL	low: ≤ 8; high: >8	equal probabilities
log removal of norovirus during wetlands subsurface flow	log_{10} units	low: ≤ 1.965; high: > 1.965	equal probabilities
norovirus concentration post wetlands subsurface flow	PCR units/mL	low: ≤ 0.09; high: > 0.09	equal probabilities
log removal of norovirus during chlorination	log_{10} units	on = 2; off = 0	equal probabilities
norovirus concentration post chlorination	PCR units/mL	low: ≤ 0.01; high: > 0.01	equal probabilities
log removal of norovirus due to spray drift control	log_{10} units	on: 1; off: 0	(NRMMC-EPHC-AHMC, 2006)
log removal norovirus due to 4 hour withholding of irrigation	log_{10} units	on: 0.4; off: 0	(Page et al., 2014)
onsite norovirus concentration	PCR units/mL	low: ≤ 0.0017; high: > 0.0.0017	equal probabilities

variable	units	states and ranges	discretisation, references[a]
wastewater volume ingested	mL	one: 1; five: 5	(Asano et al., 1992, NRMMC-EPHC-AHMC, 2006, Ryu, 2003)
dose	PCR units	low: ≤ 0.00387; high: > 0.00387	equal probabilities
risk of infection		low: $\leq 1.31 \times 10^{-5}$; high: $> 1.31 \times 10^{-5}$	equal probabilities
frequency of visits	visits per year[b]	fortnightly: 26; weekly: 52; twice weekly: 104; daily: 240	
annual risk of infection	per person per year	tolerable: $\leq 1.4 \times 10^{-3}$; high $> 1.4 \times 10^{-3}$	(Mara and Sleigh, 2010)
annual risk of illness	per person per year	tolerable: $\leq 1.1 \times 10^{-3}$; high $> 1.1 \times 10^{-3}$	(Mara and Sleigh, 2010)
annual disease burden	DALYs pppy	tolerable: $\leq 1.0 \times 10^{-6}$; high $> 1.0 \times 10^{-6}$	(World Health Organization, 2006)

[a]where indicated, thresholds and ranges for states were derived from published values

[b]for occupational exposures, 'daily' exposure is assumed to be 5 days a week for 48 weeks a year

Table E5

QMRA process model input parameters and distributions - Campylobacter

variable	units	distribution[a] or point estimates	references
Campylobacter concentration in raw wastewater	CFU/mL	triangular (0.1, 7, 100)	(Cunliffe, 2006, NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacter* during primary treatment	log_{10} units	uniform (0, 0.5)	(NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacte*r during secondary treatment	log_{10} units	uniform (1.0, 3.0)	(NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacter* during lagoon storage	log_{10} units	uniform (1.0, 5.0)	(NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacter* during wetlands surface flow	log_{10} units	1.0	(NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacte*r during wetlands subsurface flow	log_{10} units	uniform (1.0, 3.0)	(NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacter* during chlorination	log_{10} units	discrete uniform (0, 4)	(NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacter* due to spray drift control	log_{10} units	discrete uniform (0, 1)	(NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacter* due to 4 hr withholding of irrigation	log_{10} units	0.7	(Page et al., 2014)
wastewater volume ingested	mL	discrete uniform (1, 5)	(Asano et al., 1992, NRMMC-EPHC-AHMC, 2006, Ryu, 2003)
frequency of visits	visits per year	discrete uniform (26, 240)	
illness to infection ratio		0.3	(Havelaar and Melse, 2003, NRMMC-EPHC-AHMC, 2006)
disease burden	DALYs per case of illness	4.6×10^{-3}	(Havelaar and Melse, 2003, NRMMC-EPHC-AHMC, 2006)
susceptibility fraction		1	(Havelaar and Melse, 2003, NRMMC-EPHC-AHMC, 2006)

[a]distributions were defined as: triangular (minimum, most likely, maximum); uniform (minimum, maximum); discrete uniform (minimum, maximum)

Table E6

Variable states and ranges - Campylobacter

variable	units	states and ranges	discretisation, references[a]
Campylobacter concentration in raw wastewater	CFU/mL	low: ≤ 50.05; high: >50.05	equal probabilities
log removal of *Campylobacter* during primary treatment	log_{10} units	low: ≤ 0.25; high: >0.25	equal probabilities
Campylobacter concentration in primary treated wastewater	CFU/mL	low: ≤ 26.9; high: >26.9	equal probabilities
log removal of *Campylobacter* during secondary treatment	log_{10} units	low: ≤ 2.0; high: >2.0	equal probabilities
Campylobacter concentration in secondary treated wastewater	CFU/mL	low: ≤ 0.25; high: >0.25	equal probabilities
log removal of *Campylobacter* during lagoon storage	log_{10} units	low: ≤ 2.3; high: >2.3	equal probabilities
post lagoon *Campylobacter* concentration	CFU/mL	low: ≤ 2.48×10^{-4}; high: >2.48×10^{-4}	equal probabilities
log removal of *Campylobacter* during wetlands surface flow	log_{10} units	low: ≤ 0.99; high: >0.99	equal probabilities
Campylobacter concentration post wetlands surface flow	CFU/mL	low: ≤ 2.44×10^{-5}; high: >2.44×10^{-5}	equal probabilities
log removal of *Campylobacter* during wetlands subsurface flow	log_{10} units	low: ≤ 1.99; high: > 1.99	equal probabilities
Campylobacter concentration post wetlands subsurface flow	CFU/mL	low: ≤ 2.44×10^{-7}; high: > 2.44×10^{-7}	equal probabilities
log removal of *Campylobacter* during chlorination	log_{10} units	on: 4.0; off: 0	equal probabilities
Campylobacter concentration post chlorination	CFU/mL	low: ≤ 2.47×10^{-9}; high: > 2.47×10^{-9}	equal probabilities
log removal of *Campylobacter* due to spray drift control	log_{10} units	on: 1; off: 0	(NRMMC-EPHC-AHMC, 2006)
log removal of *Campylobacter* due to 4 hour	log_{10} units	on: 0.7; off: 0	(Page et al., 2014)

variable	units	states and ranges	discretisation, references[a]
withholding of irrigation			
onsite *Campylobacter* concentration	CFU/mL	low: $\leq 3.4 \times 10^{-10}$; high: $> 3.4 \times 10^{-10}$	equal probabilities
wastewater volume ingested	mL	one: 1; five: 5	(Asano et al., 1992, NRMMC-EPHC-AHMC, 2006, Ryu, 2003)
dose	CFU	low: $\leq 7.6 \times 10^{-10}$; high: $> 7.6 \times 10^{-10}$	equal probabilities
risk of infection		low: $\leq 1.5 \times 10^{-11}$; high: $> 1.5 \times 10^{-11}$	equal probabilities
frequency of visits	visits per year[b]	fortnightly: 26; weekly: 52; twice weekly: 104; daily: 240	
annual risk of infection	pppy	tolerable: $\leq 3.2 \times 10^{-4}$; high $> 3.2 \times 10^{-4}$	(World Health Organisation, 2006)
annual risk of illness	pppy	tolerable: $\leq 2.2 \times 10^{-4}$; high $> 2.2 \times 10^{-4}$	(World Health Organisation, 2006)
annual disease burden	DALYs pppy	tolerable: $\leq 1.0 \times 10^{-6}$; high $> 1.0 \times 10^{-6}$	(World Health Organization, 2006)

[a]where indicated, thresholds and ranges for states were derived from published values
[b]for occupational exposures, 'daily' exposure is assumed to be 5 days a week for 48 weeks a year

Appendix F

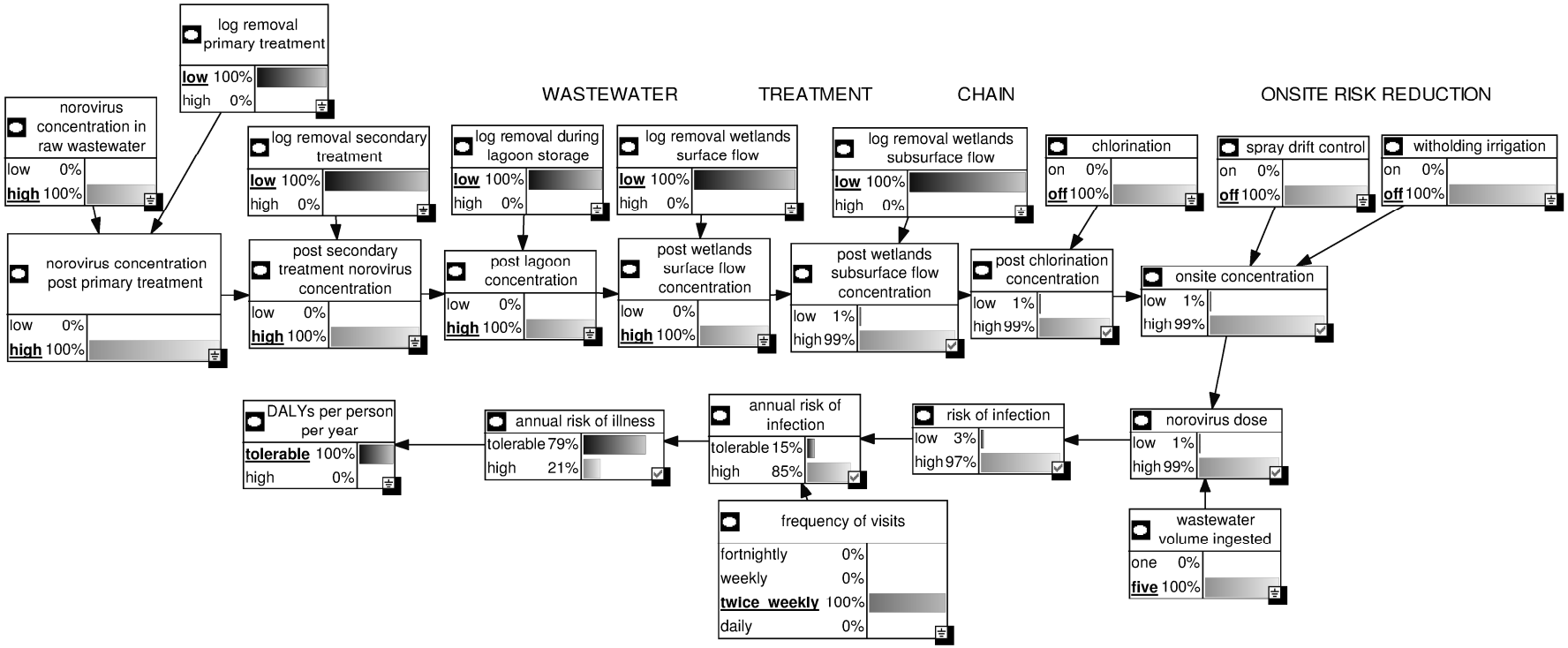

Figure F1. Scenario 2 - high infection risk conditions (norovirus) for football players, with constraint imposed of 100% tolerable DALYs per person per year.

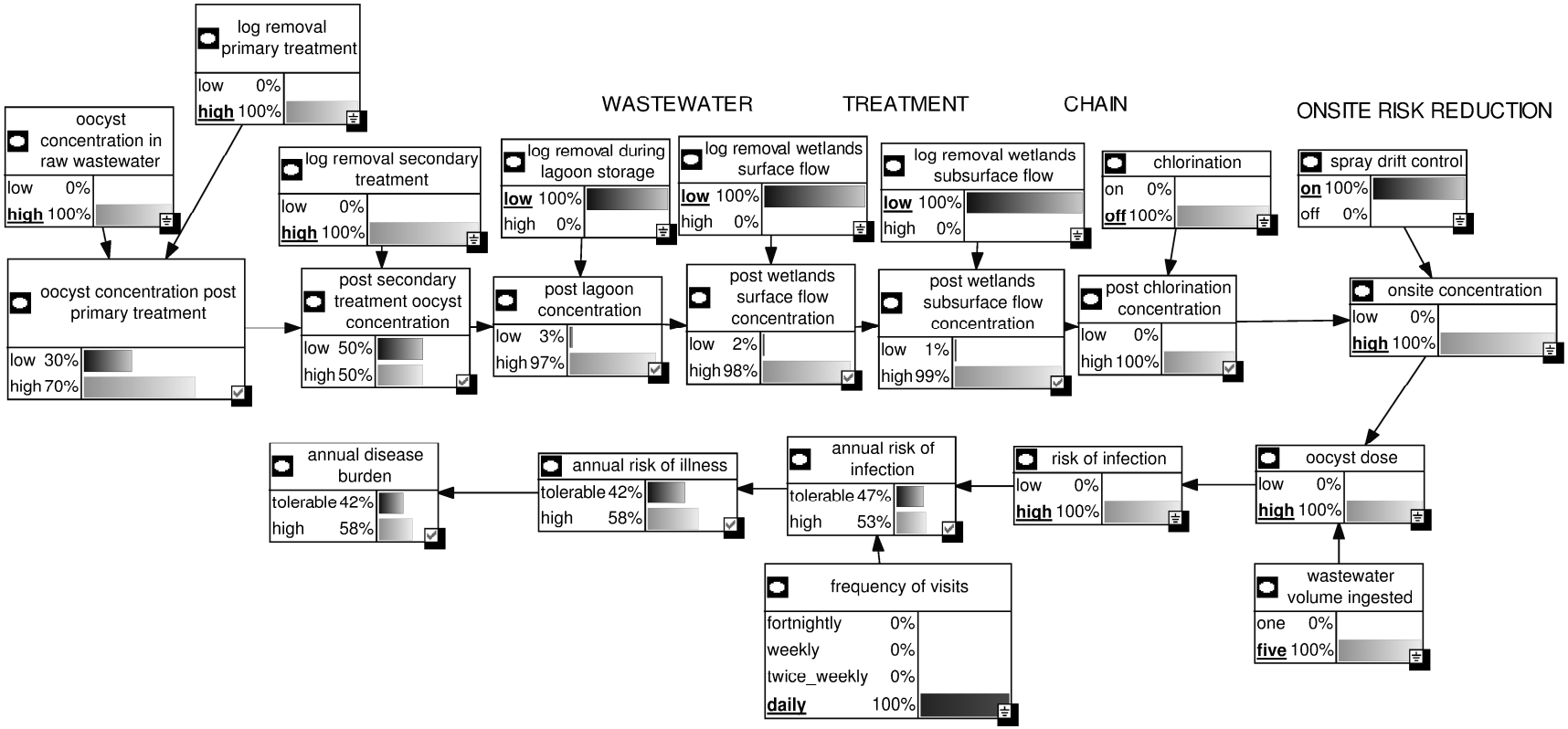

Figure F2. Scenario 3 – BN for cryptosporidiosis risk to municipal workers without chlorination.

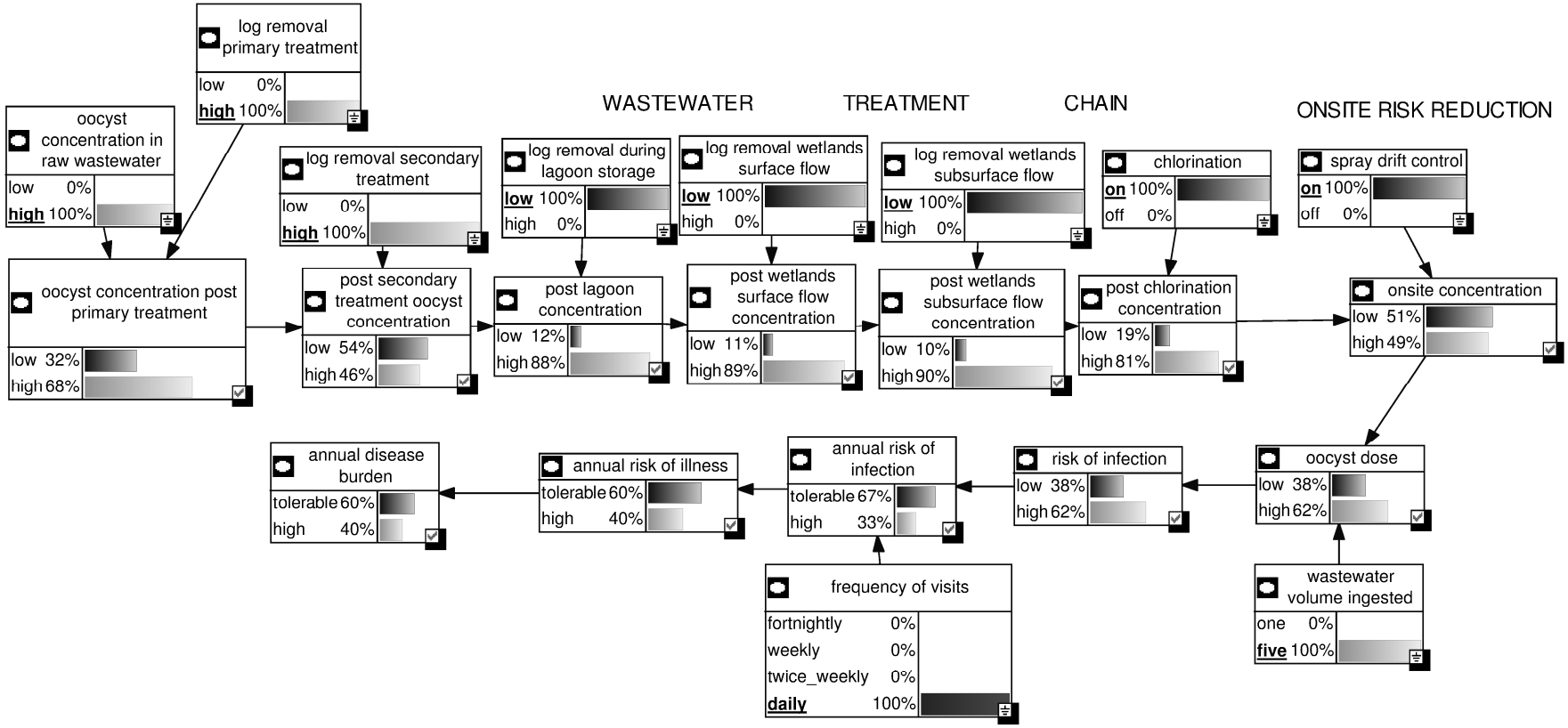

Figure F3. Scenario 3 – BN for cryptosporidiosis risk to municipal workers with chlorination.

Table F1

Scenario 3 – cryptosporidiosis risk for municipal workers: chances of response node states with and without chlorination

	annual risk of infection (%)		**annual risk of illness (%)**		**annual disease burden (%)**	
	tolerable	**high**	**tolerable**	**high**	**tolerable**	**high**
chlorination off	65	35	58	42	58	42
chlorination on	67	33	60	40	60	40
difference[a]	2	2	2	2	2	2
percent change[b]	3	6	3	5	3	5

[a]absolute value
[b]as discussed in the Method

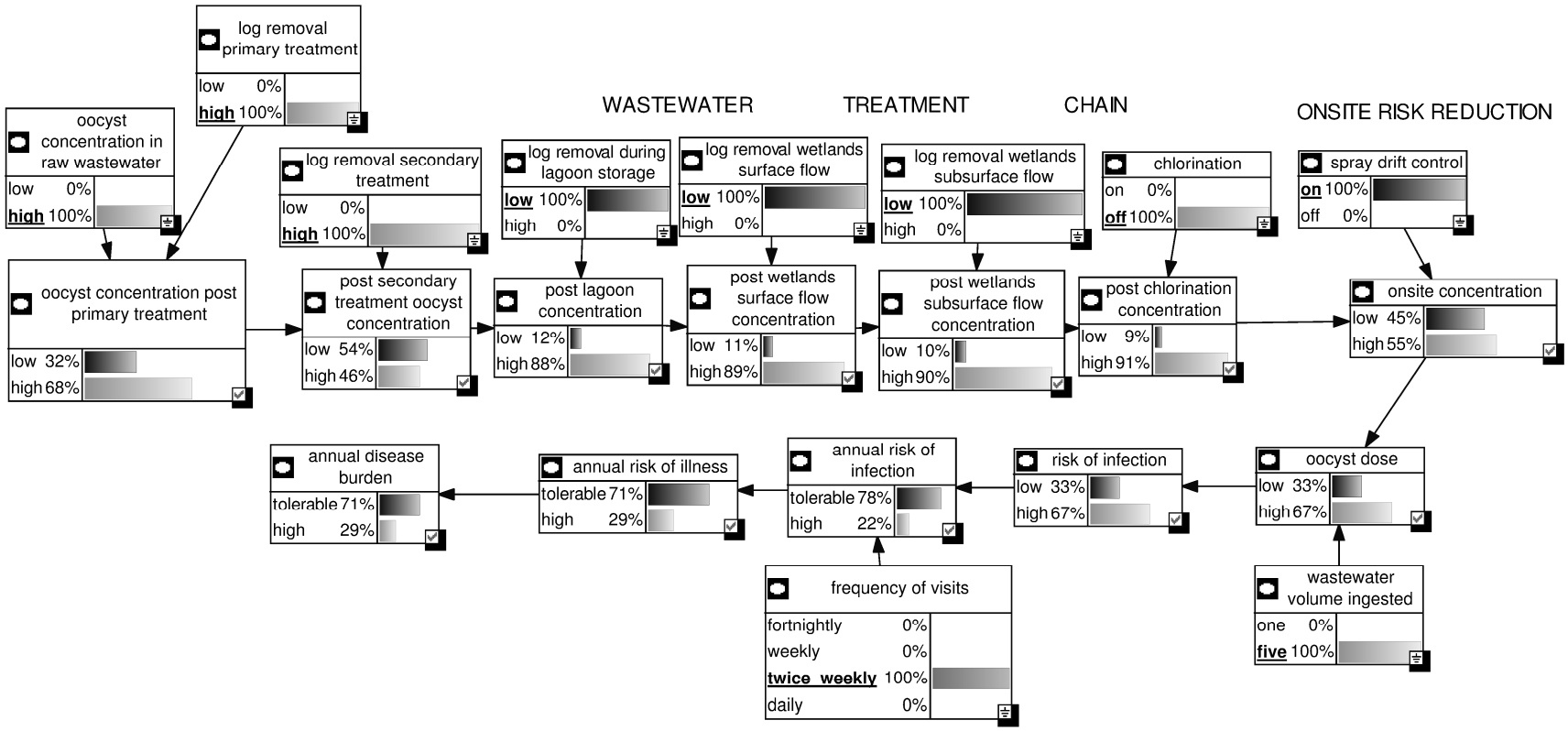

Figure F4. Scenario 3 – BN for cryptosporidiosis risk to municipal workers without chlorination, visiting twice weekly.

Table F2

Scenario 3 – cryptosporidiosis risk without chlorination, for municipal workers visiting daily and twice weekly

	annual risk of infection (%)		**annual risk of illness (%)**		**annual disease burden (%)**	
	tol.	**high**	**tol.**	**high**	**tol.**	**high**
chlorination off, visiting daily	65	35	58	42	58	42
chlorination off, visiting twice weekly	78	22	71	29	71	29
difference[a]	13	13	13	13	13	13
percent change[b]	20	37	22	31	22	31

[a]absolute value
[b]as discussed in the Method