

# Characterizing the concentration of *Cryptosporidium* in Australian surface waters for setting health-based targets for drinking water treatment

S. Petterson, D. Roser and D. Deere

## ABSTRACT

It is proposed that the next revision of the Australian Drinking Water Guidelines will include 'health-based targets', where the required level of potable water treatment quantitatively relates to the magnitude of source water pathogen concentrations. To quantify likely *Cryptosporidium* concentrations in southern Australian surface source waters, the databases for 25 metropolitan water supplies with good historical records, representing a range of catchment sizes, land use and climatic regions were mined. The distributions and uncertainty intervals for *Cryptosporidium* concentrations were characterized for each site. Then, treatment targets were quantified applying the framework recommended in the World Health Organization Guidelines for Drinking-Water Quality 2011. Based on total oocyst concentrations, and not factoring in genotype or physiological state information as it relates to infectivity for humans, the best estimates of the required level of treatment, expressed as  $\log_{10}$  reduction values, ranged among the study sites from 1.4 to 6.1  $\log_{10}$ . Challenges associated with relying on historical monitoring data for defining drinking water treatment requirements were identified. In addition, the importance of quantitative microbial risk assessment input assumptions on the quantified treatment targets was investigated, highlighting the need for selection of locally appropriate values.

**Key words** | *Cryptosporidium*, drinking water, guidelines for drinking water quality, health-based targets, quantitative microbial risk assessment

## INTRODUCTION

In line with the World Health Organization (WHO) *Guidelines for Drinking-Water Quality* (GDWQ) 2011 (WHO 2011) and the Australian *Guidelines for Water Recycling* (AGWR) (NWQMS 2006) the next revision of the Australian drinking water guidelines will include 'health-based targets'; that is, quantitative treatment targets for pathogens based on quantitative microbial risk assessment (QMRA). Specifically it is proposed that the required level of potable water treatment will be quantitatively related to the magnitude of the pathogen concentration in each source water.

One limiting factor for surface water supply treatment is expected to be protozoan contamination, in particular the presence of the 'reference' protozoan pathogen, *Cryptosporidium* due to its extreme resistance to inactivation by

chlorine (LeChevallier & Au 2004). *Cryptosporidium* has been identified in a range of hosts within Australian drinking water catchments (Cox *et al.* 2005; Power *et al.* 2005; Ryan & Power 2012) and has been identified often in high concentrations in source waters worldwide, including within Australia (Signor *et al.* 2005; Roser & Ashbolt 2007).

Information on *Cryptosporidium* concentrations in Australian surface waters is most heavily clustered around the metropolitan supply areas in south-eastern Australia where major water utilities operating in Queensland, New South Wales, Australian Capital Territory, Victoria and South Australia have been collecting data as part of their raw water monitoring programmes since the 1990s. The Water Services Association of Australia (WSAA), with the support of

**S. Petterson** (corresponding author)  
Water & Health Pty Ltd,  
PO Box 648, Salamander Bay,  
Sydney,  
NSW 2317,  
Australia  
and  
Department of Mathematical Sciences and  
Technology,  
Norwegian University of Life Sciences,  
Ås,  
Norway  
E-mail: s.petterson@waterandhealth.com.au

**D. Roser**  
Water Research Centre,  
School of Civil and Environmental Engineering,  
University of NSW,  
Sydney,  
NSW,  
Australia

**D. Deere**  
Water Futures Pty Ltd,  
Sydney,  
NSW,  
Australia

the major water utilities in these states, instigated this project to determine whether the existing historical databases held within each of these water utilities could provide the basis for estimating treatment requirements for Australian water supplies more generally. Specific objectives of the study were to:

- compile historical *Cryptosporidium* data from each project partner for several surface water sources representing a range of catchment sizes, land use and climatic conditions;
- characterize the distributions of *Cryptosporidium* in each surface water source, including the associated uncertainty in that distribution;
- quantify the treatment targets by applying the microbial 'health-based target' framework recommended in the GDWQ, and as applied in the AGWR.

## METHODS

### Data collection and compilation

For each of the six project partners, source water supplies were identified for which *Cryptosporidium* monitoring data were available, and that represented a range of different catchment types regarding the level of protection, size and land use within the catchment. For these systems the following data were requested including: *Cryptosporidium* (including total uncorrected counts), sample volume, method recovery data (i.e., method, number spiked and number recovered), results from molecular genotyping, faecal indicators (including *Escherichia coli* (*E. coli*), enterococci and spores of *Clostridium perfringens* (*C. perfringens*)) and physico/chemical data (including turbidity, total organic carbon (TOC), dissolved organic carbon (DOC), temperature and pH). All data were compiled in a common Excel datatable to aid QA/QC, e.g., checks for replicates, standardization of measurement units, generate summary statistics and facilitation selection and extraction of priority (*Cryptosporidium*) data.

### Statistical analyses of *Cryptosporidium* data

The primary objective of the statistical analyses reported here was to characterize the *Cryptosporidium* concentration

in surface water for each of the study systems, including the magnitude and variability of concentration. *Cryptosporidium* monitoring data consist of discrete counts, often with many zeros, and are the product of an imperfect detection method. Statistical analysis to facilitate appropriate parameter fitting and representative uncertainty analysis needs to be appropriate for these characteristics. Statistical models were constructed to:

1. quantify *Cryptosporidium* concentration for each study system based on discrete count data and accounting for method recovery;
2. estimate the probability that a confirmed oocyst was human infectious (where human infectivity was defined by the molecular tests applied by the two project partners who included genotyping assays in their monitoring programme).

### Quantifying *Cryptosporidium* concentrations

The models for characterizing the *Cryptosporidium* concentration were constructed based on work published by Teunis and co-workers (Teunis *et al.* 1997, 1999; Teunis & Havelaar 1999), and subsequently widely applied to QMRA for drinking water systems (Medema *et al.* 2006; Petterson *et al.* 2007; Schijven *et al.* 2011). Details are given in Appendix A (available online at <http://www.iwaponline.com/jwh/013/282.pdf>). For each system the deviance ( $-2 \times \log\text{-likelihood}$ ) was used with a likelihood ratio test to compare the fit of constant concentration model with the variable (gamma distribution) concentration model. If the difference in the deviance was greater than 3.84 ( $\chi^2$ , 95% level with 1 degree of freedom), then the addition of a second parameter to describe variable concentration was supported by the data. If the deviance was less than 3.84, it was concluded that there was no evidence in the data that the *Cryptosporidium* was variable, and the point estimate of the Poisson parameter was adequate to define source water concentrations.

### Prediction of the proportion of human infectious *Cryptosporidium*

Two project partners provided data related to the human infectivity of isolated oocysts. For these systems, all confirmed oocyst counts were pooled and analysed to

determine whether the sample contained the target genetic sequence. Statistically, a negative result was interpreted to mean that the sample contained zero human infectious oocysts; whereas a positive result indicated that the sample contained one or more human infectious oocysts.

The probability that an oocyst was human infectious was assumed to be a binomial process, with probability of human infectivity given by  $p$ . The probability of identifying  $x$  human infectious oocysts, given a total of  $n$  oocysts in the sample, is therefore given by:

$$f(x|n, p) = (1 - p)^{n-x} p^x \binom{n}{x} \quad (1)$$

For data consisting of  $m$  counts with paired infectivity result  $\delta$  ( $\delta = 1$  for positive human infectivity and  $\delta = 0$  for negative result) the likelihood function for  $p$  given an infectivity observation of  $\delta$  from  $n$  pooled oocysts is given by:

$$L(p|n, \delta) = \prod_{i=1}^m f(0|n_i, p)^{1-\delta} \cdot (1 - f(0|n_i, p))^{\delta} \quad (2)$$

For each data set, the log-likelihood function was used to find the maximum likelihood estimate (MLE) for  $p$ .

### Uncertainty analysis of parameter values

To explore the uncertainty associated with the predicted parameter values ( $\mu$ ,  $\lambda$ ,  $\rho$ ,  $p$ ), a Markov Chain Monte Carlo (MCMC) procedure was applied. For a detailed explanation of the MCMC techniques and applications see Gilks *et al.* (1996) and Gelman *et al.* (2014). Briefly, the Metropolis-Hastings algorithm was used to obtain a sample of the posterior distribution for the modelled parameters. This sample of parameter pairs was then used to construct credible intervals for the cumulative density function (CDF) for concentration; the upper 95% credible interval on the mean and the upper 95% credible interval on the upper 95% of concentration were quantified.

### Potential predictive variables

The predicted mean *Cryptosporidium* concentration was plotted against several potential predictive variables: mean

*E. coli* concentration, mean enterococci concentration, mean *C. perfringens* spore concentration, mean turbidity, mean TOC, mean DOC, system size classification, system catchment protection classification and presence or absence of a reservoir.

### Treatment targets

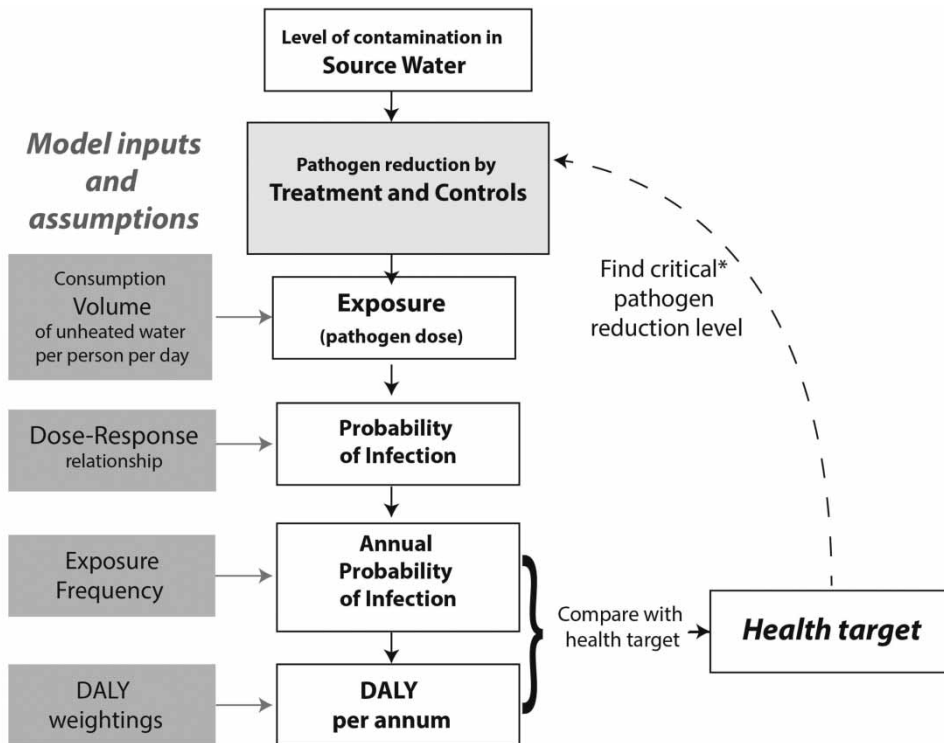
For each source water, the predicted *Cryptosporidium* concentration was used to quantify the required level of treatment in order to achieve the annual tolerable risk target of  $1 \times 10^{-6}$  disability-adjusted life year (DALY). The method applied is illustrated in Figure 1. Using the QMRA framework, for a particular inflow pathogen concentration, the level of treatment required in order to meet the health-based target was then quantified. The approach relies on fixed input assumptions for the following variables.

### Exposure volume

The total predicted exposure to pathogens depends upon the volume of unboiled water consumed per person per day. The amount of water consumed varies between individuals. In the context of setting guideline treatment limits, a reference value of 1 L per person per day is selected (consistent with GDWQ (WHO 2011)).

### Dose-response model

There are several dose-response models published for *Cryptosporidium*. The values recommended in water quality guidelines have evolved to reflect new scientific information and statistical modelling studies. Table 1 summarizes the assumptions applied in the WHO drinking water guidelines (3rd and 4th editions, e.g., WHO 2011) and the AGWR (National Water Quality Management Strategy 2006). The 3rd edition of the WHO drinking water guidelines adopted a lower value of  $r = 0.004$  based on the only results available at that time. Subsequent studies undertaken for the so-called TAMU and UCP *Cryptosporidium* isolates (and analysed using a hierarchical model by Messner *et al.* (2001)) indicated that the infectivity may be at least 10-fold higher. Further work undertaken combining these results with those for the Moredun isolate (and using a different statistical



\*The critical pathogen reduction level is the  $\text{Log}_{10}$  reduction that yields a measure of risk equal to the health target

**Figure 1** | Illustration of the WHO recommended approach for quantifying health-based targets using QMRA.

**Table 1** | Assumed *Cryptosporidium* infectivity applied in relevant water quality guidelines

Guideline document	Probability of infection per oocyst ( $r$ )	Cited study	Dose-response data
WHO GDWQ 3rd edition, 2003	0.004	Not given	IOWA isolate propagated in calves (DuPont <i>et al.</i> 1995)
AGWR (NWQMS 2006)	0.059	Messner <i>et al.</i> (2001)	Above study plus TAMU and UCP isolates (Okhuysen <i>et al.</i> 1999)
WHO GDWQ 4th edition, (WHO 2011)	0.2	Medema <i>et al.</i> (2009)	Above studies plus Moreduin isolate (Okhuysen <i>et al.</i> 2002)

modelling approach as reported in Medema *et al.* (2009)) suggested a further 10-fold increase in oocyst infectivity.

To investigate the sensitivity of the quantified treatment targets to the assumed oocyst infectivity, the calculations were undertaken separately using the assumed value of the AGWR (NWQMS 2006) and the WHO GDWQ 4th edition (WHO 2011).

Exposure frequency and annualized risk: consumption of unboiled tap water is assumed to occur every day of the year. The probability of one or more infections per year ( $P_{\text{inf}_{\text{annual}}}$ ) is given by:

$$P_{\text{inf}_{\text{annual}}} = 1 - (1 - P_{\text{inf}})^{365} \quad (3)$$

This equation is suitable for quantifying the mean annual probability of infection, or for a single point value probability of infection. When the daily probability of infection is variable, and not a constant value, it is necessary to quantify the annual probability of infection using the following bootstrapping approach:

$$P_{\text{inf}_{\text{annual}}} = 1 - \prod_1^{365} (1 - \text{Random}[P_{\text{inf}}]) \quad (4)$$

where  $\text{Random}[P_{\text{inf}}]$  is a random sample from the distribution of  $P_{\text{inf}}$ . The full distribution of the  $P_{\text{inf}_{\text{annual}}}$  can be

estimated by repeating the simulation many thousands of times in the Monte Carlo simulation.

### Probability of illness given infection

Not all infected individuals will develop symptoms of illness. The GDWQ (WHO 2011) and AGWR (NWQMS 2006) have assumed that the probability of illness, given infection, is 0.7. This value was therefore applied in the calculations.

### DALY weighting

The DALY weighting for *Cryptosporidium* applied in the GDWQ (WHO 2011) is  $1.5 \times 10^{-3}$  DALY per case of illness. The Australian DALY study (Leder *et al.* 2012) suggested a more suitable DALY weighting for Australia would be  $2.46 \times 10^{-3}$  DALY per case of illness. To investigate the sensitivity of the quantified treatment targets to the assumed DALY weight, the calculations were undertaken separately using the WHO DALY weight, and the Australian DALY weight.

Using the above assumptions, the treatment targets were calculated for the predicted mean *Cryptosporidium* concentration (for both IFA +ve and, when available, DAPI +ve oocysts) and the upper 95th credible quantile on the predicted mean. For those systems with variable oocyst concentrations, the treatment target for the upper 95th quantile of concentration was also quantified using both a point value approach (Equation (3)), and the more appropriate bootstrapped approach (Equation (4)).

## RESULTS

### Data collation

Raw water quality data were obtained from all six project partners across 25 study sites. The *Cryptosporidium*-related parameters obtained from each partner are summarized in Table 2, and the data sets for the study systems are summarized in Table 3. A quantification of method recovery is essential for the concentration prediction (Petterson *et al.* 2007). For those systems that could not provide recovery data, a conservative value of 30% was assumed. In some cases, it was not clear whether *Cryptosporidium* counts were IFA +ve or DAPI-confirmed counts. In those cases they were assumed to be DAPI-confirmed. Two project partners reported genotyping of DAPI-confirmed oocysts; however, the specific methods used, and the implications regarding the *Cryptosporidium* species and genotypes identified by these methods, are not known.

### Statistical analyses

#### Quantifying *Cryptosporidium* concentration

The results of the statistical analysis including maximum likelihood model parameters for the Poisson (no variability, Equation (A3)) statistical model and the gamma distribution (Equation (A7)) statistical models for IFA +ve oocysts are

**Table 2** | Details of *Cryptosporidium* raw data obtained from project partners

Project partner	Volumes <sup>a</sup>	Recovery	IFA +ve counts	DAPI counts	Genotyping
1	For some systems volume provided otherwise assumed to be 10 L	For some systems	Yes	For some systems	No
2	For some systems volume provided otherwise assumed to be 10 L	Yes	In some cases not clear if IFA +ve or DAPI-confirmed	In some cases not clear if IFA +ve or DAPI-confirmed	No
3	Yes	Yes	Yes	Yes	No
4	Assumed to be 10 L unless indicated otherwise	Yes	Yes	Yes	Yes
5	Assumed to be 10 L unless indicated otherwise	Yes	Yes	Yes	Yes
6	No	No	Yes	No	No

<sup>a</sup>The reported volume was assumed to be the volume tested, as opposed to the volume sampled. In discussions with the laboratory, the most common sample volume was 10 L. In addition, a review of reported concentrations implied what the sample volume had been.



**Table 3** | Summary of *Cryptosporidium* data for study systems

System code	Reservoir	Catchment protection	Scale	<i>n</i>	Total volume analysed (L)	Date range	Recovery % mean (range)	Total number of IFA-positive oocysts detected	Total number of DAPI-confirmed oocysts
1A	Reservoir	Partial	Large	83	825	4/16/1998 5/25/2011	None	2	Not done
1B	River/stream	Limited	Large	32	314	3/3/2009 5/3/2012	37.9 (12–67)	2	2
1C	Reservoir	Partial	Intermediate	25	250	3/3/2009 6/4/2011	None	4	Not done
1D	River/stream	Limited	Large	37	636	3/4/2009 7/2/2012	36.7 (5–66)	15	11
1E	River/stream	Partial	Large	75	693	25/11/1998 25/5/2011	None	16	16
2A	River/stream	Partial	Large	167	1623	1/16/2002 10/12/2011	60.8 (11–98)	14	14
2B	River/stream	Partial	Large	185	1850	2/6/2002 10/12/2011	57.99 (10–95)	45	45
2C	River/stream	Partial	Intermediate	167	5550	1/14/2002 10/21/2011	60.86 (11–98)	90	90
2D	Reservoir	Full	Large	329	4887	7/3/2000 2/20/2008	35.9 (7–91)	5	5
3A	Reservoir	Full	Intermediate	184	17927	10/7/2003 8/1/2011	42 (30–81)	4	4
3B	Reservoir	Partial	Large	104	10913	9/30/2002 2/6/2011	44.7 (6–101)	1	1
3C	River/stream	Partial	Large	371	3958	7/4/2005 7/12/2011	36.5 (5–92)	23	23
4A	Reservoir	Full	Large	62	620	11/14/2005 4/4/2011	41 (16–72)	16	2
4B	River/stream	Full	Intermediate	68	680	11/29/2004 4/4/2011	42.25 (11–70)	51	31
4C	River/stream	Partial	Large	74	740	11/29/2004 4/6/2011	39.5 (10–76)	422	284
4D	River/stream	Partial	Intermediate	94	940	9/29/1999 4/4/2011	39 (10–72)	77	32
4Ei	River/stream	Full	Intermediate	58	580	3/16/2005 4/4/2011	40.8 (10–75)	105	58
4Eii	River/stream	Partial	Intermediate	66	660	3/16/2005 4/4/2011	39 (11–74)	64	45
4Eiii	River/stream	Partial	Intermediate	75	930	8/24/2005 4/4/2011	40 (12–74)	312	173
4Eiv	Reservoir	Partial	Large	69	690	8/25/2004 6/11/2008	39 (11–73)	7	0
5Ai	River/stream	Limited	Small	138	1380	6/5/2000 6/27/2011	52 (10–98)	1361	821
5Aii	River/stream	Limited	Small	86	860	9/6/2004 7/15/2009	36 (10–77)	441	315
5Aiii	Reservoir	Limited	Intermediate	186	1860	6/5/2000 6/29/2011	50 (10–85)	42	21
5B	River/stream	Limited	Intermediate	188	1871	6/7/2000 7/29/2009	47 (10–100)	817	491
5C	River/stream	Limited	Small	262	2545	6/6/2000 7/7/2011	50 (10–100)	9928	4932

included in Table 4. The mean (expected value) of the *Cryptosporidium* concentration ( $\mu$  for the Poisson model,  $\rho\lambda$  for the gamma model including the upper 95th quantile of uncertainty on this mean from the uncertainty sample generated using MCMC simulation) and the upper 95th quantile of variability of *Cryptosporidium* concentration (also including the upper 95th quantile of uncertainty on this upper 95th quantile from the uncertainty sample generated using MCMC simulation), are also reported. The same analysis was also undertaken for DAPI-confirmed oocysts (results not shown).

The IFA +ve oocysts data set (Table 4), for systems 1A, 1B, 1C, 3A and 3B, the decrease in deviance achieved by the gamma distribution model was less than 3.84 ( $\chi^2_{95,1}$ ) and

hence the additional parameter could not be justified. For these systems, the data set did not support the hypothesis that the *Cryptosporidium* concentration was variable, and the Poisson distribution was adequate to describe the data. For the remaining systems the assumption of a variable *Cryptosporidium* concentration was statistically justified by the data set, and CDFs were constructed. Two examples are illustrated in Figure 2. In general, the uncertainty interval about the predicted distribution of *Cryptosporidium* reduced with increasing data set size, and with increasing concentration.

Uncertainty analysis for some systems produced unstable results. For system 1D, the uncertainty analysis

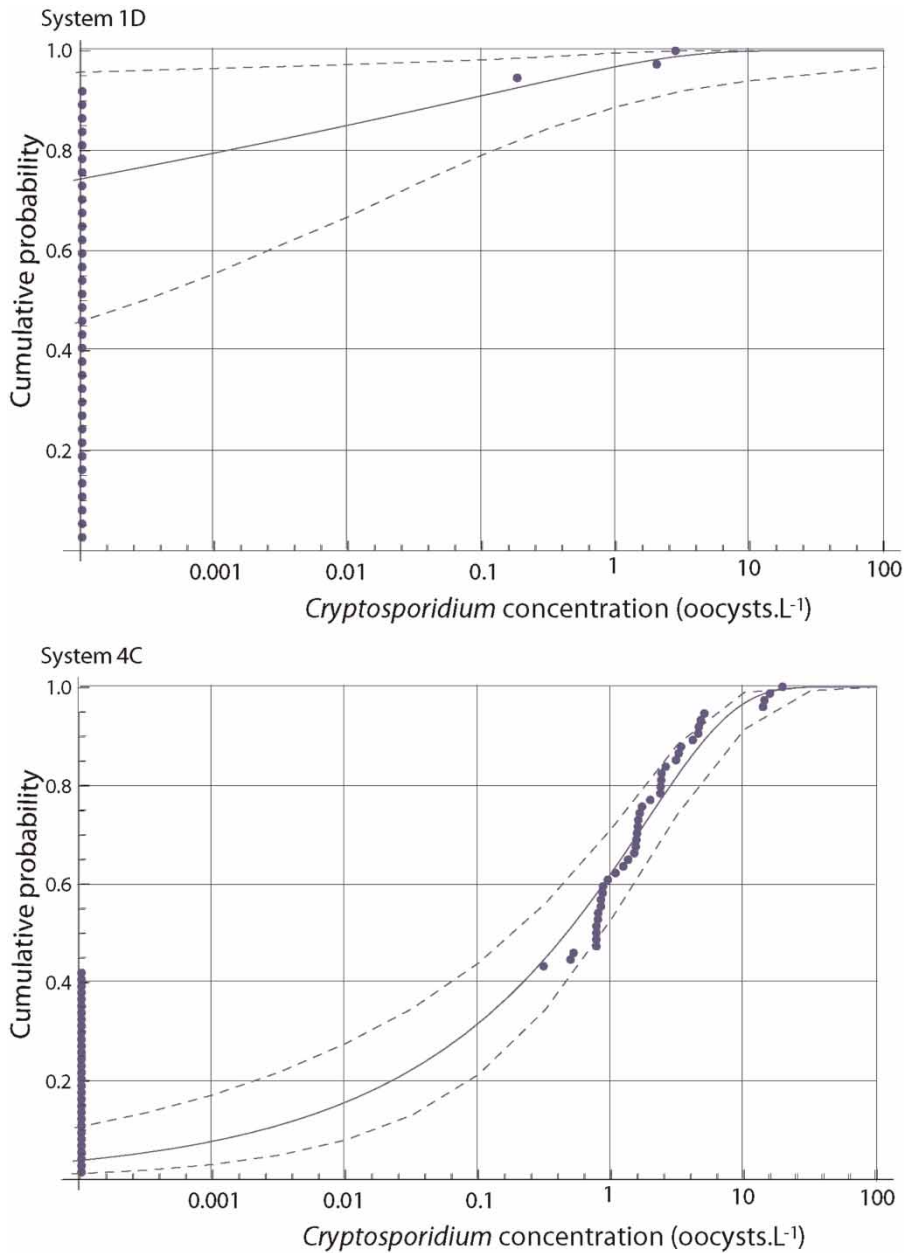
**Table 4** | Results for statistical model fitted to *Cryptosporidium* IFA +ve counts (MLE for model parameters and predicted mean (with upper 95th quantile) and upper 95th quantile on variability (with upper 95th quantile))

System code	Number of IFA-positive oocysts detected	Poisson model fit		Gamma model fit			Mean		Upper 95th quantile (oocysts L <sup>-1</sup> ) variability	
		D*	μ	D*	ρ	λ	Expected mean	Upper 95th quantile of uncertainty on mean	Expected upper 95th quantile	Upper 95th quantile of uncertainty on upper 95th
							Best estimate	Uncertainty	Best estimate	Uncertainty
1A	2	18.9	8.10 × 10 <sup>-3</sup>	18.9	5.91 × 10 <sup>-4</sup>	1.37 × 10 <sup>-7</sup>	8.10 × 10 <sup>-3</sup>	0.023		
1B	2	14.9	0.017	14.9	1.57 × 10 <sup>4</sup>	1.06 × 10 <sup>-6</sup>	0.017	0.050		
1C	4	24.1	0.053	22.9	0.31	0.17	0.053	0.12		
1D	15	119.8	0.055	34.5	0.029	4.16	0.12	145.26	1.36	377.85
1E	16	107.9	0.077	56.1	0.038	1.91	0.073	1.47	0.84	8.87
2A	14	105.0	0.014	80.3	0.061	0.22	0.013	0.035	0.15	0.38
2B	45	266.6	0.042	202.0	0.12	0.40	0.046	0.083	0.43	0.85
2C	90	617.0	0.035	374.6	0.056	0.57	0.032	0.051	0.36	0.57
2D	5	50.3	2.80 × 10 <sup>-3</sup>	38.6	0.018	0.13	2.38 × 10 <sup>-3</sup>	0.018	0.021	0.063
3A	4	46.2	5.18 × 10 <sup>-4</sup>	45.8	0.043	0.02	5.18 × 10 <sup>-4</sup>	1.17 × 10 <sup>-3</sup>		
3B	1	12.4	1.99 × 10 <sup>-4</sup>	12.3	0.040	5.90 × 10 <sup>-5</sup>	1.99 × 10 <sup>-4</sup>	7.63 × 10 <sup>-4</sup>		
3C	23	175.6	0.016				0.016	0.023		
4A	16	103.1	0.062	48.9	0.032	1.85	0.059	2.15	0.67	12.35
4B	51	246.3	0.18	138.7	0.10	2.66	0.25	0.81	2.52	8.41
4C	422	946.3	1.44	398.0	0.31	5.84	1.80	2.82	11.27	19.00
4D	77	373.5	0.21	170.2	0.07	3.47	0.25	0.65	2.70	7.03
4Ei	105	466.2	0.44	186.9	0.14	5.69	0.77	1.97	6.78	18.40
4Eii	64	288.2	0.25	132.1	0.12	1.96	0.23	0.63	2.13	6.59
4Eiii	312	965.0	0.74	307.2	0.21	4.35	0.91	1.70	6.74	13.25
4Eiv	7	58.8	0.026	28.1	0.013	2.48	0.032		0.22	
5Ai	1361	3262.2	1.90	872.9	0.38	5.73	2.18	2.98	12.41	17.18
5Aii	441	1353.2	1.42	396.2	0.25	6.01	1.48	2.42	10.24	17.42
5Aiii	42	266.9	0.046	166.2	0.069	0.63	0.044	0.092	0.47	1.00
5B	817	2653.1	1.09	786.3	0.25	3.33	0.84	1.16	5.74	8.22
5C	9928	30331.2	7.77	2066.0	0.23	41.02	9.60	12.30	67.83	87.46

\*Deviance: -2Log likelihood.

was highly unstable, evidenced by very broad uncertainty intervals (Figure 2(a)). This data set consisted of 15 IFA +ve oocysts detected in 37 samples; 13 IFA +ve oocysts were identified in one sample, with two other single detections, and the remaining samples were negative. These data suggest a highly variable concentration; however, due to the relatively small sample size and the frequency of zeros, quantifying the magnitude of that variability was

highly uncertain. For system 3C, the model fit for the gamma distribution model was unstable, and parameter estimates could not be optimized. For this data set of 371 samples, 23 oocysts were detected in 23 separate samples indicating very limited variability in concentration and hence the gamma distribution model was inappropriate. For system 4Eiv, the uncertainty analysis was unstable and did not converge on a stationary distribution for  $\rho$  and  $\lambda$ .



**Figure 2** | Examples of maximum likelihood gamma distribution CDF (solid line) and 95% credible intervals from MCMC analysis (dashed lines) fitted to *Cryptosporidium* data from (a) system 1D (above) and (b) system 4C (below) (data illustrated by points).

For this system, of 69 samples, only two samples were positive, with one of these samples containing three oocysts and the other four oocysts. This data set provided insufficient information for predicting the uncertainty in the parameter values of the gamma distribution describing variability in *Cryptosporidium* concentration.

For project partners 2 and 3, the DAPI-confirmed oocyst counts were identical to the IFA +ve. For these systems the numbers reported in Tables 5 and 6 are identical. For project partners 1, 4 and 5, the DAPI-confirmed concentrations were 0–0.08 log<sub>10</sub>; 0.12–0.95 log<sub>10</sub> and 0.13–0.29 log<sub>10</sub> lower than the IFA +ve concentrations, respectively.



**Table 5** | Maximum likelihood value of  $p$  (probability of an oocyst being human infectious) and 95% credible interval from uncertainty analysis for systems with genotyping data

System code	$p$ (%) Best estimate	95% credible interval	
		Uncertainty on $p$	
4A	0*		
4B	4.8	0.1	16
4C	1.7	0.4	4.2
4D	7.8	0.2	28
4Ei	3.1	0.1	11
4Eii	6.9	0.2	24
4Eiii	7.1	2.0	14
5Ai	25	15	38
5Aii	43	29	56
5Aiii	57	18	89
5B	31	13	54
5C	40	29	52

\*No positive results for the data set.

### Quantifying proportion of human infectious *Cryptosporidium*

The maximum likelihood estimators of the parameter  $p$  (probability of an oocyst being human infectious) for each

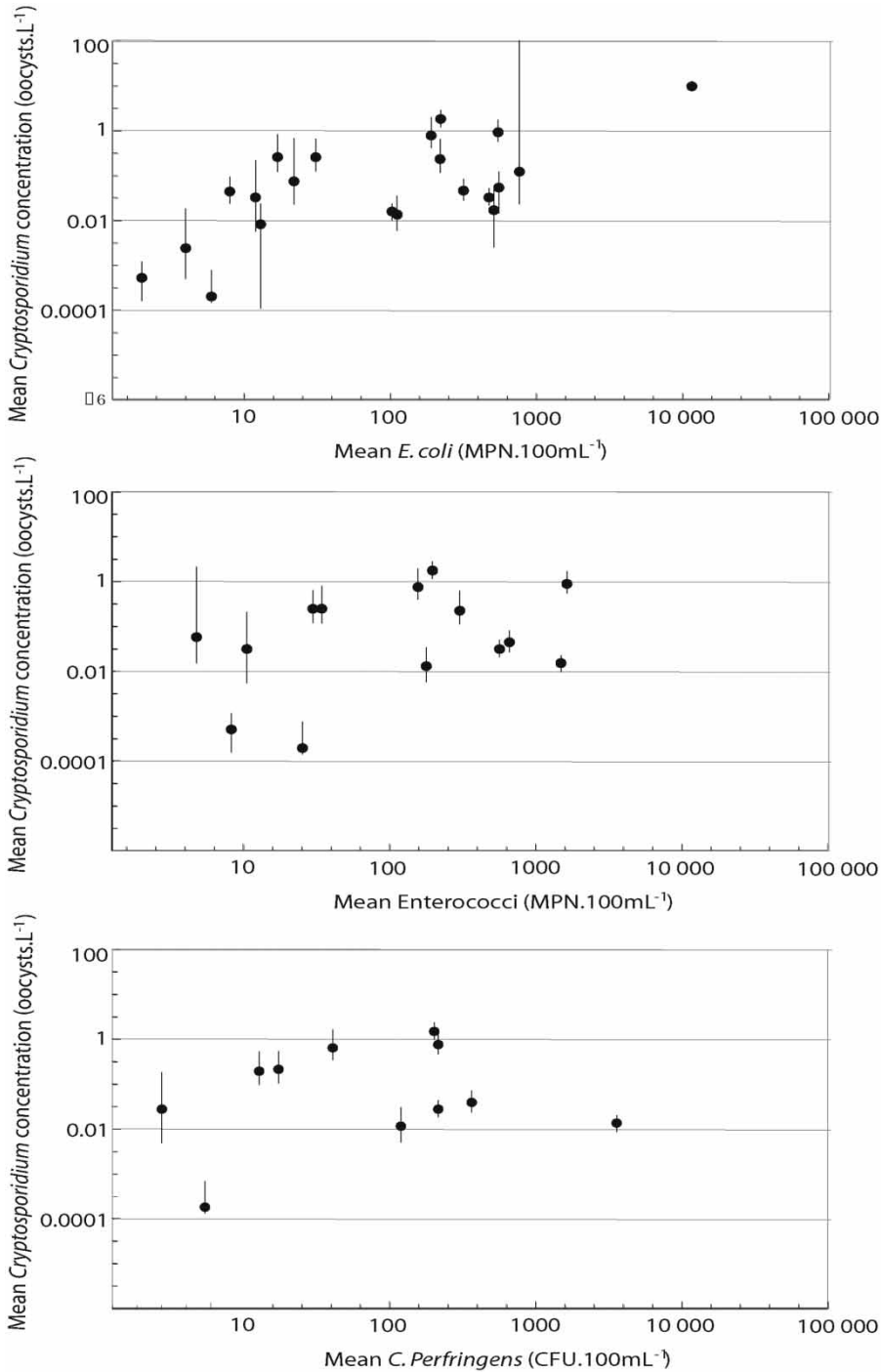
of the systems with genotyping data are summarized in Table 5. Estimated probability of human infectivity ranged from less than 2 to 57%. The probability of finding a human infectious oocyst was approximately an order of magnitude higher for systems managed by project partner 5 in comparison to project partner 4. It is essential to note that the definition of 'human infectious' is reliant upon the method used for identification. The methods used by project partners 4 and 5 are not known, and are most likely different.

### Potential predictive variables

The concentration of microbial indicators *E. coli*, enterococci and *C. perfringens* are plotted against the predicted *Cryptosporidium* (IFA +ve) concentration in Figure 3. The clearest predictive trend was with mean *E. coli* concentration. The concentrations of physicochemical parameters turbidity, TOC and DOC are plotted against predicted mean *Cryptosporidium* concentration in Figure 4. None of the three parameters indicated a clear predictive trend. The predicted mean *Cryptosporidium* concentration is plotted against the level of

**Table 6** | Comparison between annual DALY calculations using mean, upper 95th percentile (point estimate) and upper 95th percentile (bootstrapped) assuming 3 log<sub>10</sub> reduction by treatment

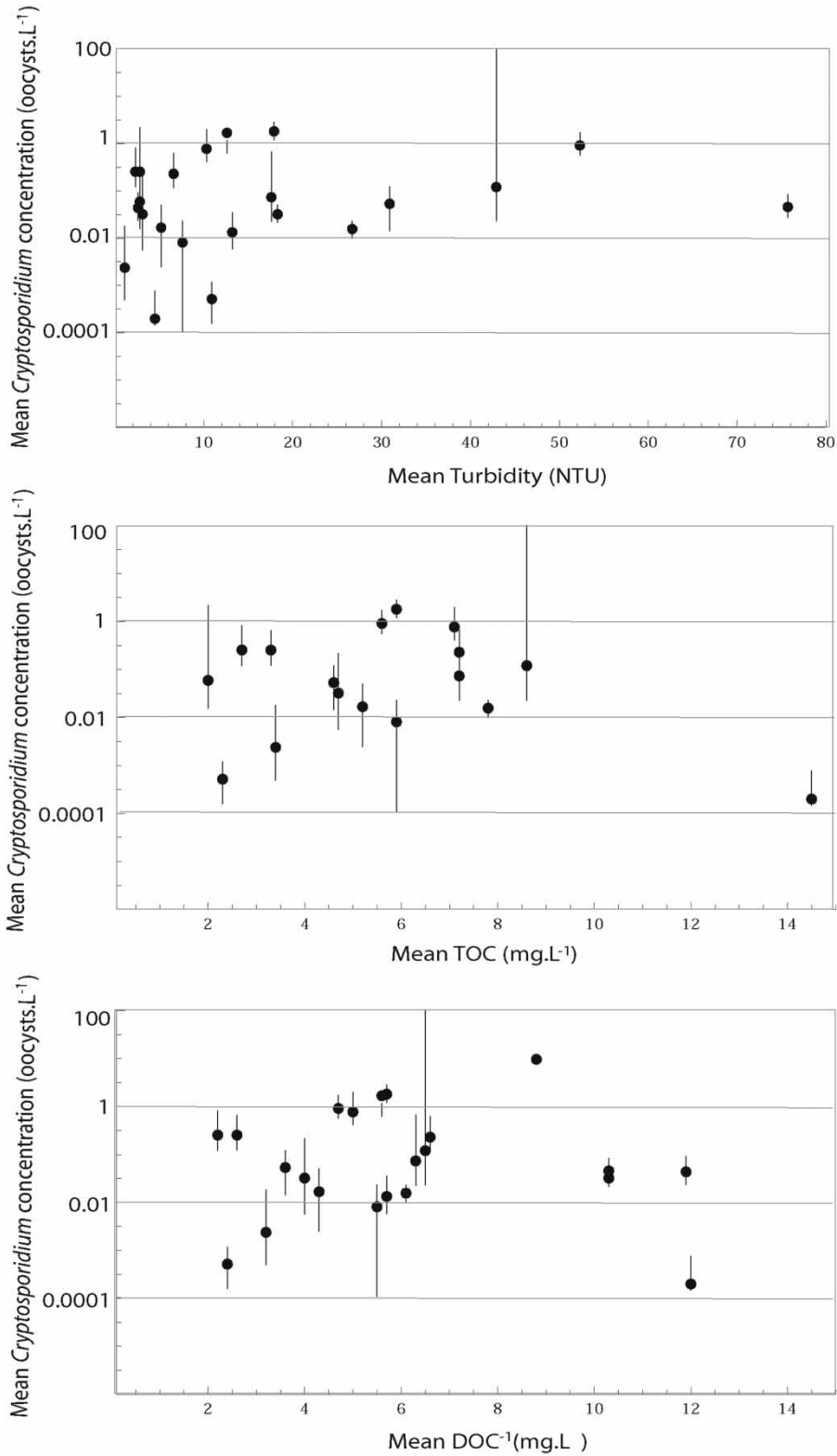
System code	Annual DALY			Log <sub>10</sub> difference between 95th percentile and mean	
	Mean	Upper 95th percentile		Point estimate	Bootstrapped
		Point estimate	Bootstrapped		
4A	$1.08 \times 10^{-6}$	$6.65 \times 10^{-6}$	$1.25 \times 10^{-6}$	0.79	0.07
4B	$5.15 \times 10^{-6}$	$5.00 \times 10^{-5}$	$7.20 \times 10^{-6}$	0.99	0.13
4C	$3.63 \times 10^{-5}$	$2.06 \times 10^{-4}$	$4.21 \times 10^{-5}$	0.75	0.08
4D	$5.13 \times 10^{-6}$	$5.34 \times 10^{-5}$	$7.32 \times 10^{-6}$	1.02	0.15
4Ei	$1.55 \times 10^{-5}$	$1.29 \times 10^{-4}$	$1.94 \times 10^{-5}$	0.92	0.11
4Eii	$4.65 \times 10^{-6}$	$4.24 \times 10^{-5}$	$6.24 \times 10^{-6}$	0.96	0.13
4Eiii	$1.83 \times 10^{-5}$	$1.28 \times 10^{-4}$	$2.25 \times 10^{-5}$	0.85	0.09
4Eiv	$6.51 \times 10^{-7}$	$4.36 \times 10^{-6}$	$1.40 \times 10^{-6}$	0.83	0.30
5Ai	$4.34 \times 10^{-5}$	$2.24 \times 10^{-4}$	$4.94 \times 10^{-5}$	0.71	0.07
5Aii	$2.95 \times 10^{-5}$	$1.88 \times 10^{-4}$	$3.57 \times 10^{-5}$	0.80	0.08
5Aiii	$8.83 \times 10^{-7}$	$9.50 \times 10^{-6}$	$1.21 \times 10^{-6}$	1.03	0.15
5B	$1.69 \times 10^{-5}$	$1.10 \times 10^{-4}$	$2.02 \times 10^{-5}$	0.81	0.08
5C	$1.78 \times 10^{-4}$	$7.67 \times 10^{-4}$	$2.19 \times 10^{-4}$	0.63	0.08



**Figure 3** | Predicted mean *Cryptosporidium* IFA +ve concentration (black dots) with 95% credible interval on the mean (lines) versus mean concentration of microbial indicators.

catchment protection, the catchment scale and the presence or absence of a reservoir in Figure 5. The mean concentration of *Cryptosporidium* was highest for

small, unprotected catchments. The presence of a reservoir led to a lower predicted mean *Cryptosporidium* concentration.



**Figure 4** | Predicted mean *Cryptosporidium* IFA +ve concentration (black dots) with 95% credible interval on the mean (lines) versus mean concentration of physiochemical parameters.

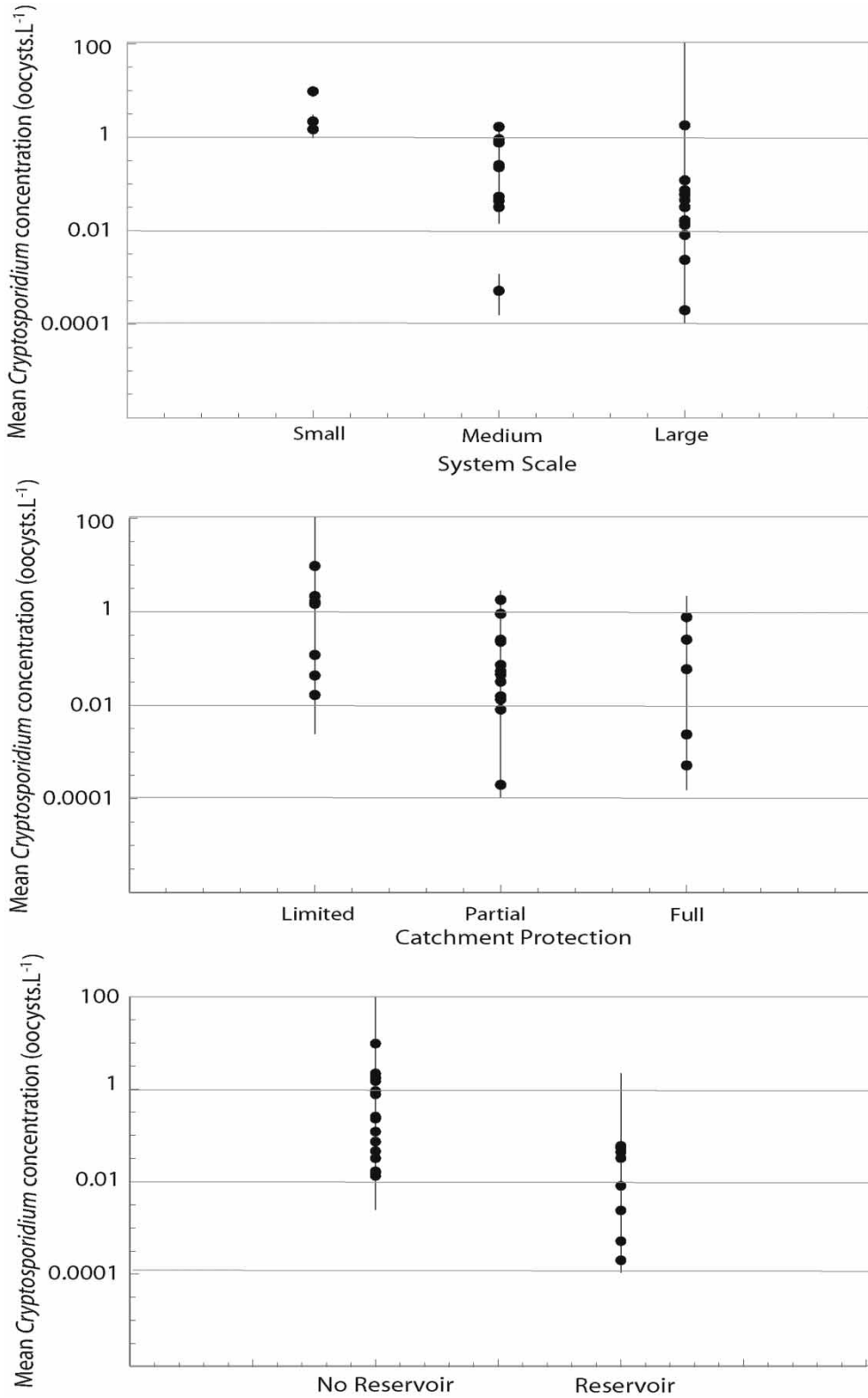


Figure 5 | Predicted mean *Cryptosporidium* IFA +ve concentration (black dots) with 95% credible interval on the mean (lines) versus catchment categories.

## Quantifying treatment targets

The impact of two input assumptions of the estimated treatment requirements the dose–response model and the DALY weighting, was tested. Comparison of the candidate dose–response parameters indicated that a change from the values applied in the AGWR (NWQMS 2006) ( $r = 0.059$ ) to GDWQ (WHO 2011) ( $r = 0.2$ ) would lead to an increase in the required level of treatment of  $0.6 \log_{10}$ . For the treatment target calculations the GDWQ (WHO 2011) assumption was applied as it was assumed to represent more recent scientific evidence, has been put together by a credible group, and has undergone extensive international peer review. Comparison of the DALY weighting assumptions demonstrated that a change from the GDWQ (WHO 2011) ( $1.5 \times 10^{-3}$ ) weighting to the Australian DALY weighting ( $2.46 \times 10^{-3}$ ) would lead to an increase in the required level of treatment of  $0.2 \log_{10}$ . For the treatment target calculations, the Australian DALY weighting was applied as it was assumed to be locally relevant.

Table 6 summarizes the results of the analysis to determine the appropriate statistic of *Cryptosporidium* concentration that should be used for quantifying the treatment targets. For a selection of systems, the table shows the annual DALY for treated water, assuming a default  $3 \log_{10}$  treatment using: the mean concentration; a point estimate of the upper 95th percentile of concentration (Equation (3)) and a bootstrapping approach (Equation (4)). When the point estimate of the upper 95th percentile was used, the quantified risk was up to  $1 \log_{10}$  higher than the mean risk. This, however, is an overestimation of the reality since the upper 95th percentile of concentration does not occur every day of the year. The bootstrapping approach accounts for the daily variability in concentration, and the difference from the mean is much lower (up to  $0.3 \log_{10}$  for the most variable system (4Eiv), and  $\approx 0.1 \log_{10}$  for other systems). The arithmetic mean was therefore considered suitably conservative and applied for the remainder of the treatment target calculations.

The treatment targets for the study systems based on IFA +ve oocyst concentration and DAPI-confirmed oocysts concentration are reported in Table 7 and illustrated in Figure 6. The best estimates for the required level of treatment to achieve the health-based target for the study systems

varied between 1.4 and  $6.1 \log_{10}$  relying on IFA +ve counts and 1.4 and  $5.8 \log_{10}$  relying on DAPI-confirmed counts.

In general, the lowest levels of treatment were required for intermediate-to-large protected systems with storage reservoirs (e.g., 3B, 3A and 2D), and the highest levels of treatment were required for small unprotected systems drawing directly from a small weir within the river or stream (e.g., 5Ai, 5Aii and 5C).

Some important exceptions were noted, for example, system 4A (a large protected system with storage reservoir and very low mean *E. coli* concentration) was classified as requiring a relatively high level of treatment at  $3.9 \log_{10}$ . This value would be reduced to  $2.9 \log_{10}$  if only DAPI-confirmed oocysts were counted. In this protected system 16 oocysts were detected from 620 L of analysed sample, of which two oocysts were DAPI-confirmed. Neither of these oocysts were positive by the genotype testing. Similarly, system 4B (a large protected system with low mean *E. coli* concentration) was classified as requiring  $4.5 \log_{10}$  treatment, which would be reduced to  $4.2 \log_{10}$  based on DAPI-confirmed oocysts. While this system is protected, 51 IFA oocysts and 31 DAPI-confirmed oocysts were detected from a total of 680 L sampled, indicating that the total *Cryptosporidium* load could not be assumed to be insignificant. Based on the genotyping test, 4.8% of these oocysts would be expected to be human infectious. If only these ‘human infectious’ oocysts were counted, the treatment target would be reduced to  $2.4 \log_{10}$ . Furthermore, some systems that are known to be impacted were classified as requiring lower levels of treatment. For example, system 1B (large unprotected system drawn directly from the river, with a high *E. coli* concentration) would require only  $3.3 \log_{10}$  as the limited data set found only two oocysts (IFA +ve and DAPI-confirmed) in their total volume sampled of 314 L.

## DISCUSSION

In this study, existing monitoring data for *Cryptosporidium* from six project partners were compiled and used to characterize the *Cryptosporidium* concentration in 25 surface water sources in Australia drawing from catchments of

**Table 7** | Treatment targets for study systems based on mean *Cryptosporidium* concentration (with upper 95th credible interval on mean)

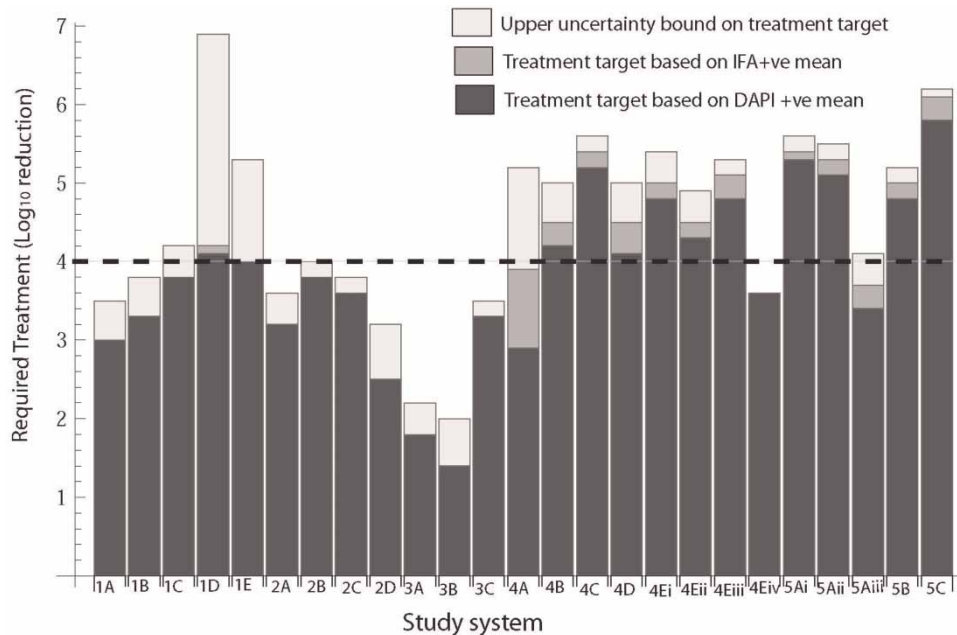
System code	Reservoir	Catchment protection	Scale	Mean <i>E. coli</i> (MPN.100 mL)	Treatment target—IFA +ve		Treatment target—DAPI +ve	
					Mean	Upper 95th credible interval on mean	Mean	Upper 95th credible interval on mean
3B	Reservoir	Partial	Large	6	1.4	2.0	1.4	2.0
3A	Reservoir	Full	Intermediate	2	1.8	2.2	1.8	2.2
2D	Reservoir	Full	Large	4	2.5	3.2	2.5	3.4
1A	Reservoir	Partial	Large	13	3.0	3.5	NA	NA
2A	River/stream	Partial	Large	112	3.2	3.6	3.2	3.6
3C	River/stream	Partial	Large	103	3.3	3.5	3.3	3.5
1B	River/stream	Limited	Large	513	3.3	3.8	NA	NA
4Eiv	Reservoir	Partial	Large	12	3.6			
2C	River/stream	Partial	Intermediate	476	3.6	3.8	3.6	3.8
5Aiii	Reservoir	Limited	Intermediate	8	3.7	4.1	3.4	4.3
2B	River/stream	Partial	Large	318	3.8	4.0	3.8	4.0
1C	Reservoir	Partial	Intermediate	555	3.8	4.2	NA	NA
4A	Reservoir	Full	Large	1	3.9	5.2	2.9	
1E	River/stream	Partial	Large	22	4.0	5.3	NA	NA
1D	River/stream	Limited	Large	768	4.2	6.9	4.1	10.3
4Eii	River/stream	Partial	Intermediate	220	4.5	4.9	4.3	5.5
4D	River/stream	Partial	Intermediate	31	4.5	5.0	4.1	4.9
4B	River/stream	Full	Intermediate	17	4.5	5.0	4.2	4.7
4Ei	River/stream	Full	Intermediate	191	5.0	5.4	4.8	5.6
5B	River/stream	Limited	Intermediate		5.0	5.2	4.8	5.0
4Eiii	River/stream	Partial	Intermediate	549	5.1	5.3	4.8	5.2
5Aii	River/stream	Limited	Small		5.3	5.5	5.1	5.4
4C	River/stream	Partial	Large	222	5.4	5.6	5.2	5.5
5Ai	River/stream	Limited	Small		5.4	5.6	5.3	5.4
5C	River/stream	Limited	Small	11,621	6.1	6.2	5.8	6.0

different sizes, land use characteristics and climatic regions. These concentrations were then used within the QMRA framework recommended by WHO (2011) for drinking water treatment, and which have already been applied to develop health-based targets for recycled water treatment in Australia via the AGWR (NWQMS 2006).

Relying on historical monitoring data gathered not for this study, but for a wide range of purposes, over many years, from a range of laboratories and using a variety of methods, presented very significant challenges when attempting to meet the purposes of the present study. First, there was no uniformity between utilities with respect to enumeration methodology and data reporting protocols.

The use of quantitative recovery controls, reporting of IFA +ve and/or DAPI – or DIC-confirmed oocysts, determination of sample volume (total volume and volume filtered), and the meaning of infectivity assays, varied both between and within utilities and was not standardized. The data supplied did not routinely come with explanations of methodology and how measurements should be interpreted. Interpretation of the data provided for the study, which was fundamental to the appropriate quantification of the *Cryptosporidium* concentration in surface water and hence any quantified treatment target, required ongoing discussion between the project team, project partners and associated laboratories, with many questions still





**Figure 6** | Treatment targets for each of the study systems based on DAPI-confirmed (DAPI +ve) mean, IFA +ve mean and the upper 95th credible interval on the predicted IFA +ve mean.

outstanding at the conclusion of the study. As a result it is strongly recommended that international guidelines for pathogen enumeration be amended to promote reporting transparency between utilities and laboratories, and to facilitate future data comparisons.

A second challenge with the interpretation of historical monitoring data was the ambiguity with respect to monitoring protocols, in particular under what conditions the samples were taken. In most cases, project partners would sample randomly (for example, every Tuesday), and may then take additional samples when triggered by specific conditions. Statistical interpretation of these data are therefore compromised as the overall data set is not truly representative of a purely random sample, nor is it representative of a clear set of defined conditions. Project partner 5 included the clearest designation within their database regarding whether a sample was considered to be 'event' affected or 'routine'. However, on further investigation, the definition of an event sample was broad and had changed over time within the organization as sampling protocols were refined. Retrospective analysis of 'event' versus 'routine' samples would therefore not represent a meaningful comparison between what, in the context of this study, would be inferred as high flow and baseline flow conditions.

An ongoing question with the development of treatment targets is the appropriateness of using the total number of IFA +ve oocysts for quantifying the *Cryptosporidium* concentration given that the method cannot distinguish between viable and non-viable oocysts, nor between human infectious and non-human infectious oocysts. From the results reported, treatment targets based on DAPI-confirmed oocysts were only  $\approx 0.2 \log_{10}$  lower than those based on IFA +ve (except for system 4A, where the difference was  $1 \log_{10}$ ). However, DAPI confirmation is a conservative marker for oocyst viability since it merely indicates the presence of intact internal contents and does not measure ability to cause an infection. In addition, while the probability of an oocyst being human infectious was quantitatively predicted in this study based on reported data, the true meaning of genotyping assays was not possible to interpret because the methods were not reported and were not consistent between project partners. The historical monitoring data therefore provided only limited scientific basis for addressing these issues of viability and human infectivity. Further, the continuing uncertainty about human infectivity and viability of *Cryptosporidium* meant that, at present, the data support much higher levels of treatment than might in fact be needed. It would be interesting to

do a cost–benefit analysis on the relative price of determining viability and infectivity versus providing an additional 1–3  $\log_{10}$  worth of treatment.

Predicting the concentration of *Cryptosporidium* in the water body from the oocyst count data required more complex statistical models than are usually applied to water quality data in order to provide the most defensible possible predictions of treatment requirements, given the above limitations. The models used in this study have been widely applied for the purpose of quantifying *Cryptosporidium* concentration for QMRA (Teunis & Havelaar 1999; Medema *et al.* 2006; Schijven *et al.* 2011) and represent current best practice. Nevertheless, as evidenced in this study, real data sets frequently provide limited information for fitting model parameters and can lead to high uncertainty intervals, particularly when pathogen concentration is low and the data sets consist of large proportions of zeros or non-detects.

Rainfall events have been linked with elevated loading of *Cryptosporidium* loading in surface water (Kistemann *et al.* 2002; Signor *et al.* 2005) and with waterborne disease (Thomas *et al.* 2006; Cann *et al.* 2013). Accounting for rainfall loading events, particularly in smaller catchments with highly variable flows, and probably quality, is an important issue for the quantification of treatment requirements that remains unaddressed. The historical data used in this study were inadequate for that purpose because the monitoring data were not collected to reflect the diversity of conditions within the study catchments. In addition, water supplies that are protected by storage reservoirs typically had lower *Cryptosporidium* concentrations and hence required lower levels of treatment. However, it may be that rainfall-induced peaks had been missed by the monitoring programme, a problem illustrated by past events at Lake Burragorang (Hawkins *et al.* 2000). These factors, and in particular how they should be incorporated within the classification of treatment targets, need further consideration.

The statistic for describing *Cryptosporidium* concentration as an input to the QMRA model was investigated by comparing the mean concentration with a point estimate of the 95th percentile, and the more appropriate bootstrapped 95th percentile of annual risk. Using a point estimate of the 95th percentile in the risk model is overly conservative for quantifying drinking water risks since it

assumes that the 95th percentile occurs every day of the year. The annual risk calculated from a bootstrapped sample ( $n = 365$ ) of daily risk demonstrated a small difference between the mean and 95th percentile ( $\approx 0.1 \log_{10}$ ). The difference between the arithmetic mean and the upper 95th percentile of the bootstrapped annual risk is reduced with increasing number of exposures per year. The AGWR uses the upper 95th percentile of pathogen concentration in sewage as a point value input to the QMRA calculations; using the mean in the drinking water guidelines would be a deviation from this national precedent. It is, however, reasonable to adopt a different approach for drinking water exposures since consumption is assumed to occur every day. In the case of wastewater recycling exposures, frequencies can vary and may only involve a single exposure event. For less frequent exposures, using the upper 95th percentile as a point value for quantifying the annualized risk is more appropriate than the mean, and is considered to be appropriately conservative.

Adoption of the QMRA approach requires reliance upon default input assumptions for water consumption, infectivity (i.e., dose–response model), illness rates and DALY weightings. The treatment targets reported in this document have been quantified using the dose–response point value assumption ( $r = 0.2$  (WHO 2011)) and the Australian DALY weighting ( $1.24 \times 10^{-3}$ , Leder *et al.* (2012)). Together, these two factors influence the treatment targets by 0.8  $\log_{10}$  in comparison to previously applied defaults in the national water recycling guidelines (AGWR) and the WHO GDWQ 3rd edition. Hence, when applying the QMRA approach for defining national treatment targets, the sensitivity of the quantified targets to the default model inputs may not be insignificant, and the scientific data most relevant to the local context selected.

While the general pattern of the treatment requirements for each system met with expectations (e.g., higher treatment required for small impacted systems, less for large protected catchments), the identified anomalies represent a concern. Misclassification leading to a significant underestimation of treatment requirements is critical to avoid but, conversely, overestimation will lead to unnecessary treatment. Based on our analysis of historical data, relying on *Cryptosporidium* monitoring data as the sole means for classifying systems and defining categories is

inadequate. Any classification system must also take into consideration catchment characteristics and land use, and the magnitude of faecal contamination. Pending research questions related to the relative persistence of oocysts in the environment, and the prevalence and identification of species and genotypes of *Cryptosporidium* of known infectivity to humans, limit the meaningfulness of quantitatively including these factors in the quantification of treatment targets.

## CONCLUSIONS

In this study, the statistical analysis and interpretation of *Cryptosporidium* data for the setting of health-based treatment targets has been undertaken based on historical monitoring data. The following recommendations ensue from the work undertaken:

- Rather than relying on historical monitoring data, a data collection programme specifically tailored for the characterization of pathogen concentrations for the purpose of setting health-based treatment requirements needs to be defined. This programme needs to consider: the number of samples, the selection of sampling conditions, the analytical methods used, and the protocols and data reporting.
- It is recommended that international reporting guidelines for pathogen enumeration data be developed to promote transparency between utilities and laboratories, and to facilitate any future data comparisons between water utilities.
- Given the limitations of monitoring data sets for capturing the full range of variability in *Cryptosporidium* loading and the importance of avoiding misclassification, treatment classifications should not be defined based on *Cryptosporidium* enumeration data alone, but also need to take into consideration catchment knowledge, sanitary surveys and *E. coli* monitoring results.
- Based on the data sets analysed in this study, the arithmetic mean is a suitable statistic for defining the *Cryptosporidium* concentration to be used for quantifying treatment targets for drinking water exposure.
- Selection of appropriate default assumptions for the QMRA model needs to be further investigated and

understood in the light of their influence upon the quantified treatment targets, and their appropriateness to the local context.

- Reasonable methods for accounting for rainfall loading events in the quantification of treatment targets are still unexplored, and need to be addressed, particularly for small flashy catchments, and water supplies protected by large storage reservoirs.

## ACKNOWLEDGEMENTS

This project was led by Water Research Australia (WaterRA) through Gareth Roeszler and Dr David Halliwell on behalf of the Water Services Association of Australia (WSAA), represented by Evelyn Rodrigues, and a consortium of water utilities led by Dr Arran Canning of Seqwater. The key investigators were the named authors of this research article as well as Prof. Peter White, School of Biotechnology and Biomolecular Sciences, University of NSW; Dr Nick O'Connor, Ecos Environmental Consulting Pty Ltd; Dr Paul Monis, Australian Water Quality Centre; A/Prof. Una Ryan, Murdoch University and Dr Martha Sinclair, Monash University. The investigators gratefully acknowledge the water utilities that provided data, information, funding and technical contributions to the project, including Seqwater, Brisbane, Queensland; Melbourne Water, Victoria; Sydney Catchment Authority, New South Wales; Water Corporation, Western Australia; South Australian Water Corporation and ACTEW, Canberra, ACT. David Sheehan is thanked for his detailed review of this manuscript on behalf of WSAA.

## REFERENCES

- Cann, K. F., Thomas, D. R., Salmon, R. L., Wyn-Jones, A. P. & Kay, D. 2013 *Extreme water-related weather events and waterborne disease*. *Epidemiol. Infect.* **141** (4), 671–686.
- Cox, P., Griffith, M., Angles, M., Deere, D. & Ferguson, C. 2005 *Concentrations of pathogens and indicators in animal feces in the Sydney watershed*. *Appl. Environ. Microbiol.* **71** (10), 5929–5934.
- DuPont, H. L., Chappell, C. L., Sterling, C. R., Okhuysen, P. C., Rose, J. B. & Jakubowski, W. 1995 *The infectivity of *Cryptosporidium parvum* in healthy volunteers*. *New Engl. J. Med.* **332** (13), 855–859.

- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. 2014 *Bayesian Data Analysis*. Chapman and Hall, Boca Raton, FL.
- Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. 1996 *Markov Chain Monte Carlo Methods in Practice*. Chapman and Hall, London.
- Hawkins, P. R., Swanson, P., Warnecke, M., Shanker, S. R. & Nicholson, C. 2000 Understanding the fate of *Cryptosporidium* and *Giardia* in storage reservoirs: a legacy of Sydney's water contamination incident. *J. Water Supply Res. Technol. AQUA*. **496**, 289–306.
- Kistemann, T., Claßen, T., Koch, C., Dangendorf, F., Fischeder, R., Gebel, J., Vacata, V. & Exner, M. 2002 Microbial load of drinking water reservoir tributaries during extreme rainfall and runoff. *Appl. Environ. Microbiol.* **68** (5), 2188–2197.
- LeChevallier, M. & Au, K. K. 2004 *Water Treatment and Pathogen Control: Process Efficiency in Achieving Safe Drinking Water*. Report of the World Health Organization, Geneva.
- Leder, K., Sinclair, M., Stevenson, C., Vos, T., Hall, R., Gibney, K., Dale, K., O'Toole, J. & Birkenstock, T. 2012 *Establishing Australian Health Based Targets for Microbial Water Quality (WQRA Project Number: 1004/09)*. <http://www.waterra.com.au/project-details/6>.
- Medema, G., Loret, J.-C., Stenström, T. A. & Ashbolt, N. 2006 *Quantitative Microbial Risk Assessment in the Water Safety Plan. Final Report on the EU MicroRisk Project*. European Commission, Brussels.
- Medema, G. J., Teunis, P. F., Blokker, M., Deere, D., Davidson, D., Charles, P. & Loret, J. F. 2009 *Risk Assessment of Cryptosporidium in Drinking Water*. World Health Organization, Geneva.
- Messner, M. J., Chappell, C. L. & Okhuysen, P. C. 2001 Risk assessment for *Cryptosporidium*: a hierarchical Bayesian analysis of human dose response data. *Water Res.* **35** (16), 3934–3940.
- National Water Quality Management Strategy (NWQM) 2006 *National Guidelines for Water Recycling: Managing Health and Environmental Risks*. Report of the National Resource Management Ministerial Council, Environment Protection and Heritage Council, Australian Health Ministers Conference.
- Okhuysen, P. C., Chappell, C. L., Crabb, J. H., Sterling, C. R. & DuPont, H. L. 1999 Virulence of three distinct *Cryptosporidium parvum* isolates for healthy adults. *J. Infect. Dis.* **180** (4), 1275–1281.
- Okhuysen, P. C., Rich, S. M., Chappell, C. L., Grimes, K. A., Widmer, G., Feng, X. & Tzipori, S. 2002 Infectivity of a *Cryptosporidium parvum* isolate of cervine origin for healthy adults and interferon-knockout mice. *J. Infect. Dis.* **185**, 1320–1325.
- Petterson, S. R., Signor, R. S. & Ashbolt, N. J. 2007 Incorporating method recovery uncertainties in stochastic estimates of raw water protozoan concentration for QMRA. *J. Water Health* **5** (Suppl. 1), 51–65.
- Power, M. L., Sangster, N. C., Slade, M. B. & Veal, D. A. 2005 Patterns of *Cryptosporidium* oocyst shedding by eastern grey kangaroos inhabiting an Australian watershed. *Appl. Environ. Microbiol.* **71** (10), 6159–6164.
- Roser, D. J. & Ashbolt, N. 2007 *Source Water Quality Assessment and the Management of Pathogens in Surface Catchments and Aquifers*. Report of the Cooperative Research Centre for Water Quality and Treatment, CRC for Water Quality and Treatment, Salisbury, SA 5108, Australia.
- Ryan, U. & Power, M. 2012 *Cryptosporidium* species in Australian wildlife and domestic animals. *Parasitology* **139** (13), 1673–1688.
- Schijven, J. F., Teunis, P. F., Rutjes, S. A., Bouwknegt, M. & de Roda Husman, A. M. 2011 QMRAspot: a tool for Quantitative Microbial Risk Assessment from surface water to potable water. *Water Res.* **45** (17), 5564–5576.
- Signor, R. S., Roser, D. J., Ashbolt, N. J. & Ball, J. E. 2005 Quantifying the impact of runoff events on microbiological contaminant concentrations entering surface drinking source waters. *J. Water Health* **3** (4), 453–468.
- Teunis, P. F. M. & Havelaar, A. H. 1999 *Cryptosporidium in drinking waters: evaluation of the ILSI/RSI quantitative risk assessment framework*. Report No. 284 550 006. Report of the National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands.
- Teunis, P. F. M., Medema, G. J., Kruidenier, L. & Havelaar, A. H. 1997 Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. *Water Res.* **31** (6), 1333–1346.
- Teunis, P. F. M., Ever, E. G. & Slob, W. 1999 Analysis of variable fractions resulting from microbial counts. *Quantitat. Microbiol.* **1**, 63–88.
- Thomas, K. M., Charron, D. F., Waltner-Toews, D., Schuster, C., Maarouf, A. R. & Holt, J. D. 2006 A role of high impact weather events in waterborne disease outbreaks in Canada, 1975–2001. *Int. J. Environ. Health Res.* **16** (3), 167–180.
- WHO 2011 *Guidelines for Drinking-Water Quality*, 4th edn. Report of the World Health Organization, Geneva.

First received 5 December 2014; accepted in revised form 19 January 2015. Available online 20 February 2015