Changes in Escherichia coli to enteric protozoa ratios in rivers

Implications for risk-based assessment of drinking water treatment requirements

Sylvestre, Émile; Dorner, Sarah; Burnet, Jean Baptiste; Smeets, Patrick; Medema, Gertjan ; Cantin, Philippe ; Villion, Manuela ; Robert, Caroline ; Ellis, Donald ; Servais, Pierre

DOI
10.1016/j.watres.2021.117707

Publication date
2021

Document Version
Final published version

Published in
Water Research

Citation (APA)
Sylvestre, É., Dorner, S., Burnet, J. B., Smeets, P., Medema, G., Cantin, P., Villion, M., Robert, C., Ellis, D., Servais, P., & Prévost, M. (2021). Changes in Escherichia coli to enteric protozoa ratios in rivers: Implications for risk-based assessment of drinking water treatment requirements. *Water Research*, *205*, 1-10. Article 117707. https://doi.org/10.1016/j.watres.2021.117707

Water Research 205 (2021) 117707

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Changes in *Escherichia coli* to enteric protozoa ratios in rivers: Implications for risk-based assessment of drinking water treatment requirements

Émile Sylvestre [a,b,*], Sarah Dorner [b], Jean-Baptiste Burnet [a,b], Patrick Smeets [c], Gertjan Medema [c,d], Philippe Cantin [e], Manuela Villion [f], Caroline Robert [e], Donald Ellis [e], Pierre Servais [g], Michèle Prévost [a]

[a] *NSERC Industrial Chair on Drinking Water, Department of Civil, Geological, and Mining Engineering, Polytechnique Montreal, Montreal, Quebec H3C 3A7, Canada*
[b] *Canada Research Chair in Source Water Protection, Department of Civil, Geological, and Mining Engineering, Polytechnique Montreal, Montreal, Quebec H3C 3A7, Canada*
[c] *KWR Water Research Institute, Groningenhaven 7, 3433 PE Nieuwegein, The Netherlands*
[d] *Sanitary Engineering, Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600GA Delft, The Netherlands*
[e] *Ministère de l'Environnement et de la Lutte contre les changements climatiques, Québec, Canada*
[f] *Centre d'expertise en analyse environnementale du Québec, Ministère de l'Environnement et de la Lutte contre les changements climatiques, Québec, Canada*
[g] *Ecology of Aquatic Systems, Université libre de Bruxelles, Brussels, Belgium*

ARTICLE INFO

Keywords:
Minimum treatment requirements
Drinking water
Quantitative microbial risk assessment
Escherichia coli
Cryptosporidium
Giardia

ABSTRACT

Minimum treatment requirements are set in response to established or anticipated levels of enteric pathogens in the source water of drinking water treatment plants (DWTPs). For surface water, contamination can be determined directly by monitoring reference pathogens or indirectly by measuring fecal indicators such as *Escherichia coli* (*E. coli*). In the latter case, a quantitative interpretation of *E. coli* for estimating reference pathogen concentrations could be used to define treatment requirements. This study presents the statistical analysis of paired *E. coli* and reference protozoa (*Cryptosporidium, Giardia*) data collected monthly for two years in source water from 27 DWTPs supplied by rivers in Canada. *E. coli*/*Cryptosporidium* and *E. coli*/*Giardia* ratios in source water were modeled as the ratio of two correlated lognormal variables. To evaluate the potential of *E. coli* for defining protozoa treatment requirements, risk-based critical mean protozoa concentrations in source water were determined with a reverse quantitative microbial risk assessment (QMRA) model. Model assumptions were selected to be consistent with the World Health Organization (WHO) Guidelines for drinking-water quality. The sensitivity of mean *E. coli* concentration trigger levels to identify these critical concentrations in source water was then evaluated. Results showed no proportionalities between the log of mean *E. coli* concentrations and the log of mean protozoa concentrations. *E. coli*/protozoa ratios at DWTPs supplied by small rivers in agricultural and forested areas were typically 1.0 to 2.0-log lower than at DWTPs supplied by large rivers in urban areas. The seasonal variations analysis revealed that these differences were related to low mean *E. coli* concentrations during winter in small rivers. To achieve the WHO target of 10^{-6} disability-adjusted life year (DALY) per person per year, a minimum reduction of 4.0-log of *Cryptosporidium* would be required for 20 DWTPs, and a minimum reduction of 4.0-log of *Giardia* would be needed for all DWTPs. A mean *E. coli* trigger level of 50 CFU 100 mL^{-1} would be a sensitive threshold to identify critical mean concentrations for *Cryptosporidium* but not for *Giardia*. Treatment requirements higher than 3.0-log would be needed at DWTPs with mean *E. coli* concentrations as low as 30 CFU 100 mL^{-1} for *Cryptosporidium* and 3 CFU 100 mL^{-1} for *Giardia*. Therefore, an *E. coli* trigger level would have limited value for defining health-based treatment requirements for protozoa at DWTPs supplied by small rivers in rural areas.

* Corresponding Author.
E-mail address: emile.sylvestre@polymtl.ca (É. Sylvestre).

https://doi.org/10.1016/j.watres.2021.117707
Received 13 July 2021; Received in revised form 11 September 2021; Accepted 22 September 2021
Available online 25 September 2021

1. Introduction

Concentrations of reference enteric protozoa (*Cryptosporidium, Giardia*) in source water must be adequately estimated to define health-based protozoa treatment requirements for drinking water safety (WHO 2016). However, data are not always available because of the difficulty and cost associated with analyzing waterborne protozoa in environmental samples. In these situations, fecal indicator bacteria (FIB) such as *Escherichia coli* (*E. coli*) are commonly used as indicators of pathogen occurrence. Although FIB monitoring data sets can provide important information on fluctuations of fecal contamination in source water, it is essential to keep in mind that this indicator has limitations for predicting concentrations of enteric protozoa (Wu *et al.* 2011, Health Canada 2017). Meteorological and environmental factors can have different effects on the fate and transport of indicators and pathogens in water. Moreover, indicators can originate from other sources than pathogens.

Monitoring of protozoa in raw water from drinking water treatment plants (DWTPs) is recommended in Canada (Health Canada 2017). Still, protozoa monitoring is not mandatory in most Canadian provinces (Government of Manitoba 2007, Gouvernement du Québec 2012). In the United States and Alberta, Canada, protozoa monitoring is mandatory for large community water supplies and small community water supplies when *E. coli* concentrations are low (USEPA 2010, Government of Alberta 2012). Most of these regulations rely on the assumption that drinking water sources exposed to high *E. coli* concentrations have a greater probability of protozoa occurrence, independently of their concentration (Payment and Locas 2011). However, a quantitative relationship between concentrations of *E. coli* and protozoa needs to be established to define health-based minimum treatment requirements for protozoa using quantitative microbial risk assessment (QMRA).

To determine whether *E. coli* data can be used to support the implementation of health-based treatment requirements, *E. coli* to protozoa ratios can be evaluated over a given period at multiple DWTPs supplied by different types of drinking water sources. A meta-analysis of *E. coli/Cryptosporidium* ratios in primary sources of fecal contamination suggested that *E. coli* is generally a good indicator for predicting *Cryptosporidium* occurrence for urban pollution sources (raw and treated wastewater) but not for agricultural runoff (Lalancette *et al.* 2014). In this study, *E. coli/Cryptosporidium* ratios were also evaluated at drinking water intakes from 13 DWTPs in Quebec, Canada, using data from Payment *et al.* (2000). As estimated in their meta-analysis, ratios at drinking water intakes were lower for sources influenced by agricultural runoff than those influenced by municipal sewage. The present study was designed to validate these findings using recent data collected at 27 surface DWTPs supplied by rivers dominated by urban, agricultural, or wildlife sources of fecal pollution in Quebec, Canada. The mathematical model of Lalancette *et al.* (2014) was also extended to evaluate the uncertainty associated with arithmetic mean *E. coli/Cryptosporidium* and *E. coli/Giardia* ratios.

The objectives of this study were (i) to derive the arithmetic mean ratio of two correlated lognormal distributions and use this model to characterize *E. coli/Cryptosporidium* and *E. coli/Giardia* ratios in source water for 27 DWTPs supplied by rivers; (ii) to investigate the influence of river types and seasons on the variation in the magnitude of the mean ratios; and (iii) to evaluate whether *E. coli* trigger levels provide valuable information for defining health-based treatment requirements for pathogenic protozoa at these DWTPs.

2. Materials and methods

2.1. Classification of sites

Source water supplies were anonymized and classified according to their annual mean flow rate (Table 1). Group A, Group B, and Group C represent rivers with mean flow rates of less than 20 $m^3 s^{-1}$, between 20 and 100 $m^3 s^{-1}$, and larger than 100 $m^3 s^{-1}$, respectively. The main land cover type for each catchment was determined visually with Google Earth.

2.2. Sampling and microbial detection methods

Paired samples were collected monthly over approximately two consecutive years between 2011 and 2020 at each of the 27 DWTPs. For the detection of *Cryptosporidium* and *Giardia*, a raw water volume varying from 10 to 60 litres was filtered on-site with an Envirochek HV cartridge (Pall) according to U.S. EPA method 1623 (from 2011 to 2012)

Table 1
Summary of *Cryptosporidium* and *Giardia* data and catchment information for 27 surface drinking water treatment plants (DWTPs).

DWTP	Main land cover type	Catchment size (km^2)	Mean river discharge (m^3/s)	*n*	*Crypto* (+ve)	*Giardia* (+ve)	Sampling period	
A01	Agricultural	100	<20	21	7	21	2018/4/30	2019/3/28
A02	Agricultural	200	<20	21	13	21	2018/4/30	2019/3/28
A1	Forested	100	<20	20	20	20	2016/5/17	2018/1/31
A2	Mixed	<100	<20	20	15	20	2016/5/10	2018/1/31
A3	Mixed	500	<20	21	12	21	2014/6/17	2016/3/21
A4	Agricultural	<100	<20	24	18	24	2014/3/25	2016/3/15
B1	Mixed	2500	23	22	19	22	2014/3/25	2016/3/15
B2	Forested	4000	26	19	14	19	2016/5/10	2017/11/6
B3	Mixed	2500	26	18	7	18	2016/05/9	2017/11/13
B4	Mixed	4200	27	15	6	15	2011/8/28	2013/8/13
B5	Mixed	1100	36	18	13	17	2016/5/17	2017/11/13
B6	Mixed	2500	70	18	9	18	2016/5/10	2017/11/6
B7	Agricultural	3400	74	16	9	16	2011/5/3	2013/9/23
C1	Mixed	10,000	114	19	16	18	2016/5/9	2017/11/7
C2	Agricultural	10,000	114	17	14	17	2016/5/9	2017/11/7
C3	Mixed	7000	114	15	14	15	2014/3/25	2016/9/8
C4	Mixed	10,000	190	22	8	22	2014/3/25	2016/3/15
C5	Agricultural	10,000	190	15	5	15	2011/5/3	2013/9/23
C6	Urban	>50,000	286	48	20	43	2013/1/1	2016/12/31
C7	Urban	>50,000	286	16	8	16	2011/8/22	2013/9/10
C8	Mixed	23,000	330	17	10	17	2014/3/25	2016/9/22
C9	Mixed	23,000	330	15	3	15	2014/6/17	2015/9/22
C10	Urban	>50,000	1365	45	13	39	2013/1/1	2016/12/31
C11	Urban	>50,000	1365	46	13	41	2013/1/1	2016/12/31
C12	Urban	>50,000	16,000	16	10	16	2011/5/2	2013/9/10
C13	Mixed	>50,000	16,000	16	11	16	2011/8/22	2013/9/10
C14	Mixed	>50,000	16,000	17	6	17	2011/5/2	2013/9/23

and U.S. EPA method 1623.1 (from 2013 to 2020). For one DWTP, 10-liter cubitainers were collected on-site and filtered at the laboratory. The enumeration of Cryptosporidium oocysts and Giardia cysts was carried out according to U.S. EPA method 1623 or 1623.1 (USEPA 2005a, (USEPA 2012)). Mean analytical recovery rates of 43 matrix spikes in raw water samples collected at these DWTPs were 0.46 (Standard Deviation [SD] = 0.14) for *Cryptosporidium* and 0.50 (SD = 0.17) for *Giardia*.

The enumeration of *E. coli* was done by membrane filtration according to the modified mTEC method from 2011 to 2012 or mFC-BCIG method (CEAEQ MA.700 - Ec.BCIG 1.0) from 2013 to 2020. These methods involve filtration of the sample through a 0.45 µm porosity membrane, which is then deposited on a selective agar medium and incubated at 44.5°C for 24 h. The *E. coli* colonies were identified and counted visually, and the concentration was expressed in CFU 100 mL^{-1}.

3. Statistical methods

3.1. Ratio of microbial concentrations

3.1.1. Poisson counts in mixture distributions

A pragmatic way to account for non-detects is to assume that each observed count is Poisson distributed and that the unknown microbial concentration is described by a mixture distribution (Haas *et al.* 1999). Within a mixed Poisson modeling framework, the probability of finding k organisms in a homogenous sample x of volume V collected from a suspension of mean concentration c is given by a Poisson distribution with probability mass function:

$$p(k|cV) = \frac{(cV)^k}{k!}\exp(-cV) \tag{1}$$

Overdispersion relative to the Poisson distribution was expected because microbial concentrations typically vary over orders of magnitude in river water. A mixing distribution was selected to account for the unobserved heterogeneity, i.e., the temporal variation of the concentration c in successive samples. The mixed Poisson distribution describing variations in counts is given by:

$$P(k|cV,\theta) = \int^{p} (k;cV) f(c;\theta) dc \tag{2}$$

where f is a mixing distribution of parameters θ. The precise form of the mixed Poisson distribution depends on the choice of $f(c;\theta)$. In this study, the lognormal distribution was selected as a mixing distribution because the mean ratio of two lognormal variables is simple to derive. The lognormal distribution has a probability density:

$$f(c|\mu,\ \sigma) = \frac{1}{\sigma c\sqrt{2\pi}}\exp\left[-\frac{1}{2}\frac{[ln\ c-\mu]^2}{\sigma^2}\right] \tag{3}$$

where the shape parameter μ and the scale parameter σ are, respectively, the expected value and the standard deviation of the natural logarithm of c.

Original observations (counts, processed volumes) were not available to model temporal variations in *E. coli* concentrations. Therefore, a parameters of the lognormal distribution were estimated from reported *E. coli* concentrations. Non-detects were replaced by a detection limit of 1 CFU 100 mL^{-1}. This approach for handling non-detects should not result in substantial estimation biases because the proportion of non-detects was typically very small (< 5%) for *E. coli*. Recovery rates of 100% were assumed for *E. coli*.

3.1.2. Ratio of two correlated lognormal variables

To derive the ratio of lognormal variables, let X be the *E. coli* concentration in CFU L^{-1} and Y be the protozoa concentration in (oo)cyst L^{-1}. If $X \sim LN(\mu_X,\ \sigma_X)$ and $Y \sim LN(\mu_Y,\ \sigma_Y)$, then, by definition, $\ln(X)$ and $\ln(Y)$ are normally distributed. It follows that the difference Z between $\ln(X)$ and $\ln(Y)$ is normally distributed with mean

$$\mu_Z = \mu_X - \mu_Y \tag{4}$$

and variance

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 - 2\sigma_n^{xy} \tag{5}$$

where σ_n^{xy} is the covariance between X and Y in log space. Therefore, the ratio X/Y is described by $Z \sim LN(\mu_Z,\ \sigma_Z)$ with mean

$$E(Z) = \exp\left(\mu_Z + \frac{\sigma_Z^2}{2}\right). \tag{6}$$

To account for the correlation between X and Y, the covariance σ_n^{xy} in Eq. (5) can be calculated as follows:

$$\sigma_n^{xy} = \ln\left(\frac{\sigma_{ln}^{xy}}{\bar{x}\bar{y}} + 1\right) \tag{7}$$

where σ_{ln}^{xy} is the covariance in real space and $\bar{x}$ and $\bar{y}$ are the arithmetic means of X and Y (Crow and Shimizu 1987). The covariance σ_n^{xy} is related to the Pearson' moment correlation coefficient in log-space ρ_n^{xy} by:

$$\rho_n^{xy} = \frac{\sigma_n^{xy}}{\sigma_X\sigma_Y} \tag{8}$$

By substituting Eqs. (7) – (8) in Eq. (5), the mean ratio X/Y from Eq. (6) can be rewritten in the following form:

$$E(X/Y) = \exp\left((\mu_X - \mu_Y) + \frac{\sigma_X^2 + \sigma_Y^2 - 2\rho_n^{xy}\sigma_X\sigma_Y}{2}\right) \tag{9}$$

The covariance σ_{ln}^{xy} and sample means $\bar{x}$ and $\bar{y}$ in Eq. (7) were evaluated from the sample considering that non-detects had a concentration of 0 (oo)cyst. The arithmetic mean ratio was evaluated because it can be calculated straightforwardly. It is also a standard measure of location to compare ratios among groups. Recovery rates of 100% were assumed for protozoa in this model.

3.1.3. Influence of seasonality, weather events and nonconstant analytical recovery

Mean ratios of pooled data by seasons were evaluated for Group A, Group B and Group C to provide an overall summary of the influence of river types and seasons on the magnitude of mean ratios. The climate in southern Quebec is humid continental and it is characterized by large seasonal temperature differences and four seasons. In this work, seasons were defined as winter (Dec.-Feb.), spring (Mar.-May), summer (June-Aug.) and autumn (Sept.-Nov.).

For DWTPs A4, C6 and C7, mean ratios in routine monitoring conditions (i.e., ratios evaluated with routine monitoring data) were compared to daily mean ratios determined for event conditions (i.e., ratios evaluated with data obtained following rainfall and snowmelt events). A detailed characterization of event-based sampling strategies implemented to obtain these data sets is presented elsewhere (Sylvestre *et al.* 2020b, Sylvestre *et al.* 2021). Daily mean ratios in event conditions were estimated by averaging ratios from each sample i as follows:

$$E\left(\frac{X}{Y}\right)_{event} = \frac{1}{n}\sum_{i=1}^{n}\frac{X_i}{Y_i} \tag{10}$$

where X is the *E. coli* concentration and Y is the protozoa concentration. Event-based samples were collected at regular intervals of 4 h or 6 h over 24 h (DWTPs A4, C6) or three times over 96 h (DWTP C7). Concentrations of protozoa during events were estimated from raw (oo)cyst counts corrected for sample-specific recovery rates. To account for the analytical recovery in the estimation of the mean ratio in routine monitoring conditions, the number of microorganisms observed in each sample was modeled as a binomial distribution of independent counts having a

probability of recovery described by a beta distribution with parameters $(\widehat{\alpha}, \widehat{\beta}) = (6.48, 7.70)$ for *Cryptosporidium* and $(\widehat{\alpha}, \widehat{\beta}) = (3.80, 3.91)$ for *Giardia* (Sylvestre *et al.* 2020c).

3.1.4. Model implementation

The parameters of the Poisson lognormal distributions were estimated using Markov chain Monte Carlo (MCMC) simulations via rjags (v4–6) (Plummer 2013) in R (v4.0.4). For each parameter, four Markov chains were run for 1×10^6 iterations after a burn-in phase of 10^4 iterations. The convergence of chains was monitored using the Brooks-Gelman-Rubin scale reduction factor (Gelman and Shirley 2011). The median of the posterior distribution of the arithmetic mean and its 95% credibility interval were reported for microbial concentrations and ratios. Estimates of the 95% credibility interval of the posterior distributions of the parameter values were considered reasonably accurate when an effective sample size (ESS) higher than 10,000 was obtained (Kass *et al.* 1998, Kruschke 2014). Uninformative or weakly informative priors were adopted. Priors on parameters μ and σ were set to Uniform (-10, 10) and exp(1), respectively. The rationale for the selection of these priors is presented elsewhere (Sylvestre *et al.* 2020c). The R code is provided in the Supplementary Material.

3.2. Scaling relationship between E. coli and protozoa

3.2.1. Bivariate power-law regression

Sample mean *E. coli* concentrations and sample mean protozoa concentrations from the 27 sites were compared on a log-log scale. The relationship between *E. coli* and protozoa was evaluated with a bivariate power-law expressed as follows:

$$\overline{c_2} = r\overline{c_1}^{k} \tag{11}$$

where $\overline{c_2}$ is the mean protozoa concentration, $\overline{c_1}$ is the mean *E. coli* concentration, r is the intercept at $\overline{c_1} = 1$, and k is the dimensionless scaling exponent. The slope k identifies the relationship between $\overline{c_1}$ and $\overline{c_2}$. The following three situations are relevant to consider: (1) $k = 1$, the relationship between $\overline{c_1}$ and $\overline{c_2}$ is proportional; (2) $k < 1$, a large change in $\overline{c_1}$ is only related to a small change in $\overline{c_2}$; (3) $k = 0$: a large change in $\overline{c_1}$ is not related to a change in $\overline{c_2}$. Power laws were fitted using the iteratively reweighted least-squares (IRLS) algorithm from the glm function in R software environment.

3.2.2. Risk-based critical mean Cryptosporidium and Giardia concentrations

Risk-based critical mean *Cryptosporidium* and *Giardia* concentrations in source water were determined for treatment reduction requirements of 3.0-log and 4.0-log. The critical mean concentration was defined as the maximum *Cryptosporidium* or *Giardia* concentration that would be tolerable to achieve a health-based target of 10^{-6} disability-adjusted life year (DALY) per person per year. Critical mean concentrations C were calculated with the following reverse QMRA model:

$$\mathrm{C} = \left[\frac{1}{R}\cdot \mathrm{I}\cdot 10^{-T}\cdot \mathrm{V}\cdot r\cdot P_{ill|\mathrm{inf}}\cdot \mathrm{DB}\cdot \mathrm{E}\cdot f_s\right]^{-1}\cdot \mathrm{HT} \tag{11}$$

where R is the recovery rate of the enumeration method, I is the infectivity of (oo)cysts, T is the reduction of the organism by treatment, V is the volume of drinking water ingested per person per day, r is the probability of infection for a single organism, $P_{ill|inf}$ is the probability of infection given illness, DB is the disease burden, E is the number of exposures per year, f_s is the susceptible fraction of the population, and HT is the health outcome target. Each variable of the model was described by its arithmetic mean. Mean values and assumptions for each parameter are presented in Table 2. Full-scale reductions of *Giardia* and *Cryptosporidium* at the selected DWTPs were not evaluated and considered in the QMRA. Risk-based concentrations were determined for theoretical treatment reduction requirements. Most assumptions for the QMRA for *Cryptosporidium* (ingestion volume, dose-response model for infection, probability of illness given infection, disease burden) were selected to be consistent with those recommended in the WHO Guidelines for drinking-water quality (GDWQ) (WHO 2017). Conservative default mean infectivity fractions of 30% for *Cryptosporidium* oocysts and 100% for *Giardia* cysts were assumed to characterize source water concentrations. Conservative human pathogenicity fractions of 100% were assumed for *Cryptosporidium* and *Giardia*.

Table 2
Assumptions and mean values for each parameter of reverse QMRA models used for the determination of risk-based critical mean *Cryptosporidium* and *Giardia* concentrations.

Parameter	Units	*Cryptosporidium*	*Giardia*
Health outcome target (HT)	DALY per year	1×10^{-6}	1×10^{-6}
Susceptible fraction (f_s)	Fraction of population	1.00	1.00
Disease burden (DB) [a]	DALY per case	1.5×10^{-3}	1.7×10^{-3}
Risk of illness given infection ($P_{ill\|inf}$) [b]	Probability of illness per infection	0.70	0.40
Number of exposures per year (E)	Days	365	365
Risk of infection (r) [c]	Probability of infection per organism	0.20	0.02
Consumption of unheated drinking water (V) [d]	Litres per day	1	1
Treatment (T)	Log_{10}	3.0-log or 4.0-log	3.0-log or 4.0-log
Infectivity of organisms (I) [e]	Fraction of organisms	0.30	1.00
Recovery rate of the enumeration method (R) [f]	Fraction of organisms	1.00	1.00

[a] Disease burden for *Cryptosporidium* from Havelaar and Melse (2003) and for *Giardia* from Gibney *et al.* (2014)
[b] Fraction for *Cryptosporidium* from Teunis *et al.* (2002) and for *Giardia* from Nash *et al.* (1987)
[c] Dose–response for *Cryptosporidium* from WHO (2009) and for *Giardia* from Regli *et al.* (1991)
[d] Based on WHO (2017)
[e] Conservative assumption for *Cryptosporidium* based on Lalancette *et al.* (2012)
[f] Recovery rates of 100% were assumed because observations were not corrected for the analytical recovery.

3.2.3. Sensitivity of E. coli for identifying critical mean protozoa concentrations

The sensitivity of *E. coli* trigger levels to identify "high" *Cryptosporidium* and *Giardia* concentrations in source water was evaluated based on the methodology presented in USEPA (2005b). *Cryptosporidium* and *Giardia* "levels of concern" were defined as the risk-based critical mean concentrations for treatment requirement of 3.0-log. Selected *E. coli* trigger values ranged from 5 CFU 100 mL^{-1} to 100 CFU 100 mL^{-1}. The sensitivity of *E. coli* was defined as the proportion of DWTPs with "high" *Cryptosporidium* and *Giardia* concentrations that exceed the *E. coli* trigger level.

4. Results

Site-specific mean *E. coli*/protozoa ratios were typically 1.0 to 2.0-log lower at DWTPs supplied by small and medium rivers (Group A, Group B) in comparison with DWTPs supplied by large rivers (Group C) (Fig. 1). Ratios varied by approximately 4.0-log for both *Cryptosporidium* (10^3–10^7) and *Giardia* (10^0–10^4). Mean *E. coli*/*Cryptosporidium* ratios

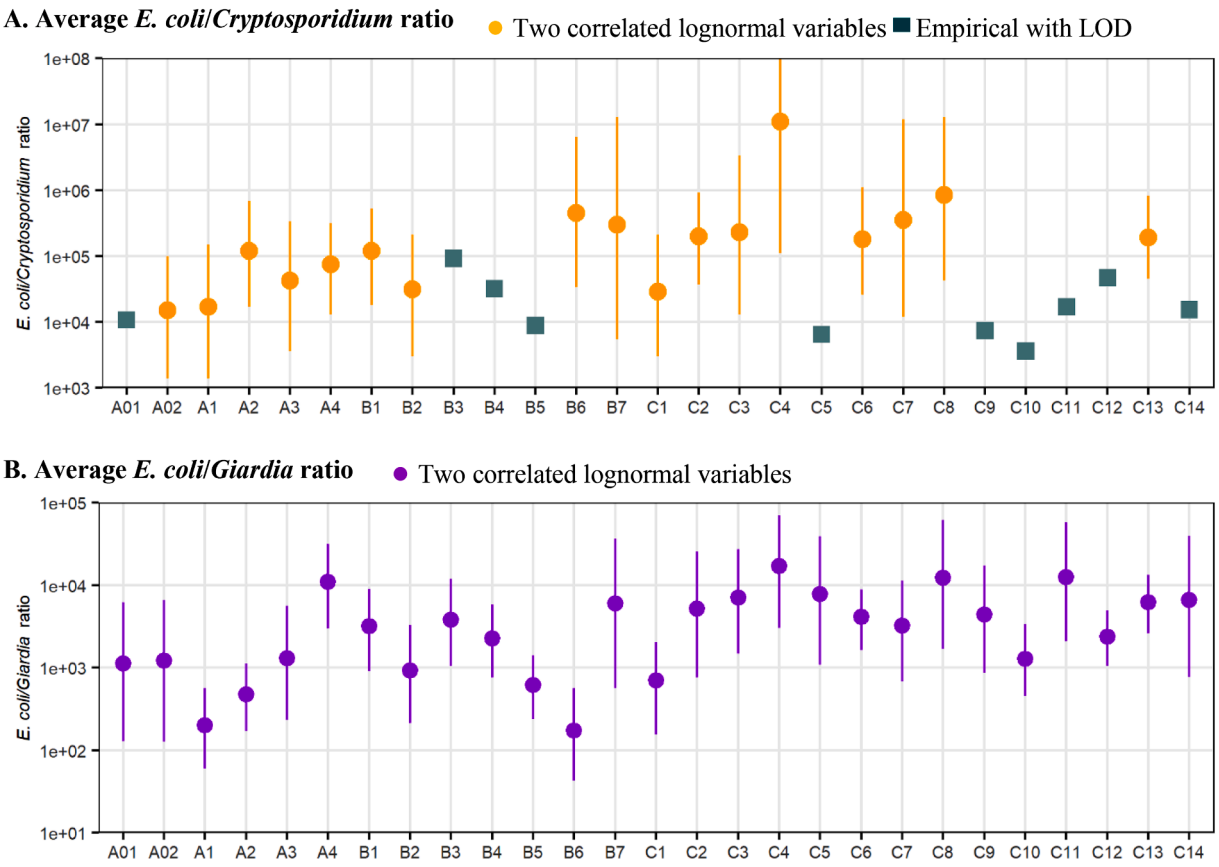

Fig. 1. *E. coli/Cryptosporidium* and *E. coli/Giardia* ratios in raw water from 27 drinking water treatment plants evaluated from paired samples collected monthly for approximately two years. Vertical bars represent the 95% credible interval on the arithmetic mean. Modeled *E. coli/Cryptosporidium* ratios are not presented for DWTPs A01, B3, B4, B5, C5, C9, C10, C11, C12 and C14 because model fits were not acceptable. Empirical ratios calculated assuming a limit of detection of 1 organism/analyzed volume are indicated for these sites.

were generally 2.0 to 3.0-log higher than mean *E. coli/Giardia* ratios because *Cryptosporidium* concentrations were lower and more uncertain than *Giardia* concentrations (Supplementary Fig. 1). Modeled *E. coli/Cryptosporidium* ratios were not reported for 10 DWTPs because MCMC samples for estimating parameters of the Poisson lognormal distribution for *Cryptosporidium* were highly correlated (ESS < 10,000). Empirical ratios (calculated assuming a detection limit of 1 (oo)cyst/volume) are shown for these sites. The 95% credible intervals on the mean spanned from 2.0 to 3.0-log for *E. coli/Cryptosporidium* ratios and from 1.0 to 2.0-log for *E. coli/Giardia* ratios. The parametric uncertainty on the *E. coli/Cryptosporidium* ratio was high when less than ten positive samples were measured at a site (Fig. 1, Table 1).

The incorporation of Pearson's product-moment correlations in the model had a minor effect (< 0.5-log) on mean ratios at most locations (Supplementary Table 1, Supplementary Fig. 1). However, high correlations (absolute value > 0.5) increased (negative correlation) or reduced (positive correlation) mean ratios by approximately 1.0-log. Overall, positive correlations were higher for *Giardia* than for *Cryptosporidium*.

The seasonal analysis of pooled data showed that *E. coli/Cryptosporidium* and *E. coli/Giardia* ratios were especially low during winter (Dec.-Feb.) in Group A ($\overline{Q} \leq 20$ m^3 s^{-1}) (Fig. 2, Supplementary Table 3). These low ratios were associated with low *E. coli* concentrations and high *Giardia* concentrations during this period. Mean *E. coli* concentrations in winter conditions were lower in Group A (47 CFU 100 mL^{-1}) compared to Group B (222 CFU 100 mL^{-1}) and Group C (668 CFU 100 mL^{-1}). The seasonal analysis also shows that the uncertainty on the mean *Cryptosporidium* concentration dominated the uncertainty on the *E. coli/Cryptosporidium* ratio (e.g., summer ratio in Group C).

During event conditions and routine monitoring conditions, mean *E. coli/Cryptosporidium* ratios were similar for DWTPs A4 (rainfall) and C6 (rainfall and snowmelt) (Fig. 3). In contrast, *E. coli/Giardia* ratios at DWTPs A4 and C7 (snowmelt) were approximately 1.0-log lower during peak conditions in comparison to routine monitoring conditions. Accounting for beta distributed recovery rates reduced the mean *E. coli/protozoa* ratios but did not change the magnitude of the 95% credible interval.

The power-law regression showed that the log of mean *E. coli* concentrations and the log of mean protozoa concentrations were not proportional (Fig. 4). The slope *k* for *Cryptosporidium* and *Giardia* were 0.26 (95% CI 0.08 0.47) and 0.19 (95% CI −0.04 0.43), respectively. Predicted R-squared were small for *E. coli* as a predictor of protozoa (*Cryptosporidium*: 26%; *Giardia*: 10%) and moderate for *Giardia* as a predictor for *Cryptosporidium* (43%) (Supplementary Fig. 2). To achieve a target of 10^{-6} DALY per person per year, a minimum reduction of 4.0-log of *Cryptosporidium* (critical concentration: 0.038 cysts L^{-1}) was required for 20 DWTPs. A minimum reduction of 3.0-log of *Cryptosporidium* was needed for the seven other DWTPs. For *Giardia*, a reduction of at least 4.0 log of *Giardia* (critical concentration: 0.228 cysts L^{-1}) was required for all DWTPs. Moreover, a reduction of at least 5.0 log of *Giardia* (critical concentration: 2.28 cysts L^{-1}) was required for 6 DWTPs. A treatment requirement of 3.0-log for a mean *E. coli* concentration below a trigger level of 50 CFU 100 mL^{-1} would result in a maximum risk of $10^{-5.4}$ DALY/pers.-year for *Cryptosporidium* (DWTP A1) and $10^{-4.7}$ DALY/pers.-year for *Giardia* (DWTP A1).

The sensitivity of *E. coli* trigger levels for correctly identifying "high" *Cryptosporidium* and "high" *Giardia* concentrations is shown in Table 3. A trigger level of 50 CFU 100 mL^{-1} resulted in a sensitivity of 95% for

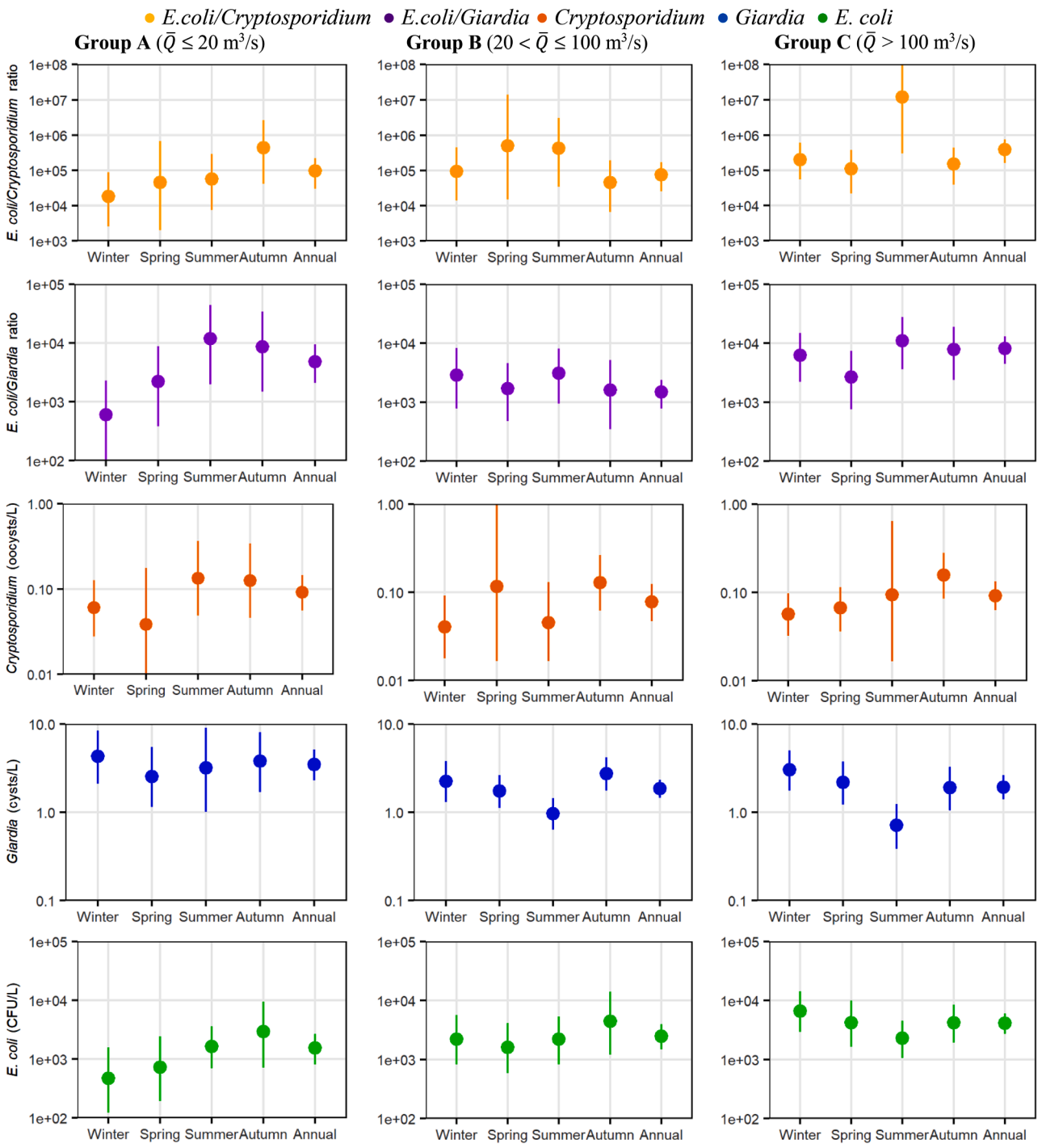

Fig. 2. Seasonal variations in *E. coli/Cryptosporidium* ratios, *E. coli/Giardia* ratios and concentrations of *Cryptosporidium, Giardia* and *E. coli* in raw water from pooled data for Groups A, B, and C. Vertical bars represent the 95% credible interval on the arithmetic mean ratio.

Cryptosporidium. Lowering the trigger level to 10 CFU 100 mL^{-1} produced a sensitivity of 100%. For *Giardia*, a trigger level of 10 CFU 100 mL^{-1} resulted in a sensitivity of 92%.

5. Discussion

Despite inherent uncertainties associated with quantifying reference pathogen concentrations from indicator data, fecal indicator bacteria trigger levels are still commonly used to define treatment requirements in drinking water safety regulations. The values and limitations of indicators for predicting reference pathogen concentrations need to be rigorously assessed to support the development and the revision of risk-based regulations. This study was undertaken to determine whether *E. coli* is a valid quantitative indicator for defining protozoa treatment requirements at surface DWTPs.

The ratio of two lognormal variables was derived and used to evaluate site-specific ratios from paired monitoring data collected at 27 DWTPs. The proposed mixed Poisson model allowed to handle non-detects, correlations, and parametric uncertainties. Mean *E.coli/Cryptosporidium* ratios predicted by this model were higher than ratios previously estimated at DWTPs supplied by rivers by Lalancette *et al.* (2014). This finding was expected because non-detects were not

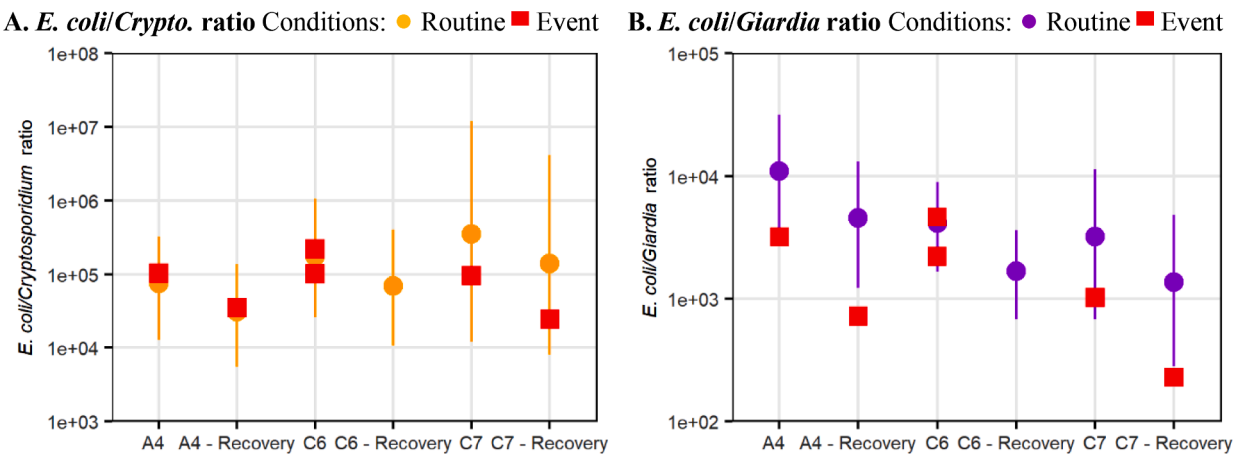

Fig. 3. *E. coli/Cryptosporidium* ratio and *E. coli/Giardia* ratio in routine monitoring conditions versus event conditions at drinking water treatment plants A4, C6, and C7. Ratios corrected for the analytical recovery of *Cryptosporidium* and *Giardia* are compared to ratios evaluated with uncorrected *Cryptosporidium* and *Giardia* concentrations. Vertical bars represent the 95% credible interval on the arithmetic mean ratio.

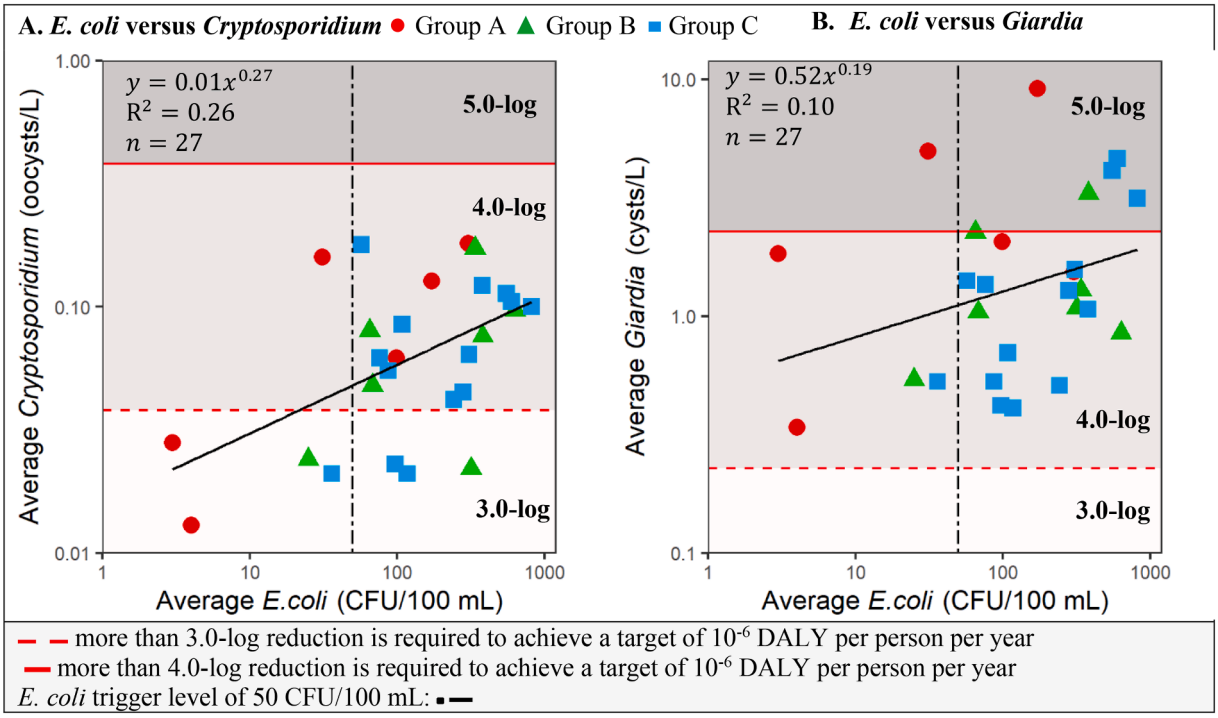

Fig. 4. Bivariate power laws predicting the relationship between the log of the sample mean *E. coli* concentrations and the log of the sample mean protozoa concentrations for 27 drinking water treatment plants. Shaded areas represent intervals for which a minimum treatment requirement of 3.0-log, 4.0-log or 5.0-log is required to achieve a target of 10^{-6} DALY per person per year. Risk-based critical concentrations were calculated with a reverse-QMRA model.

considered in ratios reported by Lalancette *et al.* (2014). Therefore, modeling ratios with mixed Poisson distributions can considerably increase the mean ratio, especially for DWTPs with a high proportion of non-detects. Accounting for correlations between concentrations of *E. coli* and protozoa rather than statistical independence did not change the magnitude of mean ratios for most DWTPs, because correlations were typically weak. However, neglecting moderately large linear correlations ($\rho > 0.5$) resulted in an overestimation or an underestimation of the mean ratio for some DWTPs, as previously demonstrated for the estimation of the product of two correlated lognormal variables (Smith *et al.* 1992). The quantification of parametric uncertainty revealed that site-specific mean *E. coli/Cryptosporidium* ratios evaluated from small data sets ($n < 30$) were often highly uncertain. Other positively skewed distributions may be used to estimate ratios. The ratio of Poisson gamma distributions could be a valuable model for datasets with small count observations and a high proportion of non-detects. This ratio is simple to derive and has been used previously for microbial risk assessment (Teunis *et al.* 2009). It can be demonstrated that if $X \sim \Gamma(\alpha_1, \beta_1)$ and $Y \sim \Gamma(\alpha_2, \beta_2)$ are independently distributed, then X/Y has a generalized beta prime distribution $\beta'\left(\alpha_1, \alpha_2, 1, \frac{\beta_2}{\beta_1}\right)$ (Johnson *et al.* 1995). However, the generalized Beta prime distribution has an infinite mean if its shape parameter α_2 has a value ≤ 1. In our work, this condition was too restrictive to model mean ratios at most DWTPs. The ratio of two correlated gamma variables has also been derived and used to evaluate the mean ratio of fecal coliforms to fecal streptococci in surface water (Loáiciga and Leipnik 2005) but the calculation of this ratio requires

Table 3
Sensitivity of *E. coli* trigger levels for correctly identifying risk-based critical mean concentrations of *Cryptosporidium* and *Giardia* at 27 drinking water treatment plants.

Protozoa	*E. coli* trigger level	10 CFU/ 100 mL		50 CFU/ 100 mL		100 CFU/ 100 mL	
	Did plant exceed *E. coli* trigger level?	No	Yes	No	Yes	No	Yes
Crypto.	≤ 0.038 oocysts/L (3.0-log required)	2	5	4	3	5	2
	> 0.038 oocysts/L (>3.0-log required)	0	20	1	19	7	13
	Sensitivity	100%		95%		65%	
Giardia	≤ 0.228 cysts/L (3.0-log required)	0	0	0	0	0	0
	> 0.228 cysts/L (>3.0-log required)	2	25	5	22	12	15
	Sensitivity	92%		81%		55%	

advanced integration techniques. Nonetheless, it has recently been demonstrated that the gamma distribution can underestimate the magnitude of *E. coli* and protozoa concentrations in source water during wet weather events (Sylvestre *et al.* 2020a, Sylvestre *et al.* 2020b). Thus, the ratio of two lognormal variables may be a better model than the ratio of two gamma variables to account for the influence of high microbial concentrations on mean ratios.

As observed by Lalancette *et al.* (2014) for the *E.coli/Cryptosporidium* ratio, *E. coli*/protozoa ratios in small and medium rivers were typically 1.0 to 2.0-log lower than those in large rivers. The analysis of pooled data with power-law regressions also showed that a large change in the mean *E. coli* concentration is related to small changes in mean *Cryptosporidium* or *Giardia* concentrations. Therefore, relying on *E. coli* data only may lead to an inaccurate assumption that protozoa concentrations are low too. The variability and uncertainty in concentrations of *E. coli* and protozoa were not considered in this analysis.

The critical mean *Cryptosporidium* concentration determined with the reverse-QMRA model for a reduction by a treatment of 3.0-log (0.038 oocysts L^{-1}) was similar to the level of 0.075 oocysts L^{-1} used to trigger additional treatment requirements (> 3.0-log) for filtered systems within the U.S. EPA Long Term 2 Enhanced Surface Water Treatment Rule (LT2) (United States Environmental Protection Agency (USEPA) 2006). In our work, a conservative dose–response model (r = 0.2) was selected for *Cryptosporidium parvum* (WHO 2009). It was assumed that this dose–response model was consistent with recent scientific evidence as it is currently applied in the WHO Guidelines for drinking-water quality (GDWQ) (WHO 2017). Nonetheless, the *Cryptosporidium* dose–response relationship remains highly uncertain at low-dose exposures (Messner and Berger 2016, Schmidt and Chappell 2016). In the current study, a default mean infectivity fraction of 30% was selected to characterize *Cryptosporidium* oocysts in source water. This fraction is similar to the one used for defining protozoa treatment requirements for surface water supplies in the United States (United States Environmental Protection Agency (USEPA) 2005). An infectivity fraction of 30% has also been recently suggested as a conservative estimate for QMRA based on *Cryptosporidium* data collected across nine locations in South Australia (Swaffer *et al.* 2018). Considering a human pathogenicity fraction of 100% for *Cryptosporidium* may overestimate risks depending on the catchment type and the inputs from wildlife. Fractions of human pathogenic species (*C. hominis* and *C. parvum*) of approximately 10% were found in the South Nation River and the Grand river basins in Ontario, Canada (Wilkes *et al.* 2013, Lapen *et al.* 2016). The analysis of site-specific human pathogenicity/infectivity data could be valuable in an in-depth QMRA to determine whether additional treatment is needed.

The *E. coli* trigger level of 50 CFU 100 mL^{-1} used to trigger *Cryptosporidium* monitoring for small community water supplies in the LT2 had a high sensitivity for identifying critical mean *Cryptosporidium* concentrations. However, increasing the trigger level to 100 CFU 100 mL^{-1} considerably reduced the sensitivity. The sensitivity of *E. coli* trigger levels was lower for *Giardia* compared to *Cryptosporidium*. Overall, these analyses showed that treatment requirements higher than 3.0-log would be needed at some sites with low mean *E. coli* concentration to achieve 10^{-6} DALY per person per year targets for *Cryptosporidium* and *Giardia*. Additional treatment credits may be obtained by optimizing the performance of existing treatment processes (e.g., filter effluent turbidity target) (USEPA 2010). However, additional disinfection processes (e.g., UV disinfection) may be needed if treatment processes are already optimized.

The evaluation of seasonal variations of ratios pooled by river sizes generally reduced the parametric uncertainty on mean ratios. The seasonal variations analysis indicated that differences in ratios were partly related to differences in mean *E. coli* concentrations during winter among groups. In fact, decreases in mean *E. coli* concentrations during winter and spring were observed at DWTPs supplied by small rivers in agricultural and forested areas but not at DWTPs supplied by large rivers in urban areas. The die-off of *E. coli* in winter may be higher in non-point source pollution (surface runoff, soil leaching) than in point source pollution (wastewater treatment plants effluents, combined sewer overflow discharges). Large rivers in southern Quebec are primarily influenced by treated and untreated municipal wastewater discharges (combined sewer overflow discharges, sporadic sewage by-passes) (Payment *et al.* 2000) and the increase *E. coli* concentrations following wastewater discharges in winter and spring has been documented previously for DWTPs C6, C7, C10, and C11 (Madoux-Humery *et al.* 2016, Burnet *et al.* 2019, Sylvestre *et al.* 2021).

By contrast, the contribution of agricultural pollution sources may be considerably higher in summer than in winter. In summer, grazing surfaces are contaminated by livestock feces and manure is applied to croplands. Consequently, rainfall-induced runoff can carry water contaminated by fecal wastes into surface water. On the other hand, in winter, livestock is kept inside due to cold temperatures; therefore, runoff water is less contaminated by livestock feces. This hypothesis is only valid for agricultural areas (not for forested areas); however, it has been shown that in rural areas, fecal contamination of rivers mainly originates from grazing areas (George *et al.* 2004, Garcia-Armisen and Servais 2007, Ouattara *et al.* 2011). Other potential causes for seasonal variations in rural areas are the buffering of surface runoff by snow in winter and the effect of freeze–thaw cycles on the survival of microorganisms in water. Wang *et al.* (2019) recently reported a decline in initial *E. coli* concentrations of 4.0-log in river water after 1 to 3 freeze–thaw cycles but not at a constant temperature of 4 °C (< 1.0-log). Afolabi *et al.* (2020) inoculated animal feces into water and measured high reductions in initial *E. coli* concentrations after one freeze–thaw for red deer feces (1.0 to 3.0-log) but not in dairy cow feces (< 1.0-log). The effect of freeze–thaw cycles on *Cryptosporidium* and *Giardia* survival is uncertain. Kato *et al.* (2002) reported a reduction of 1.0-log of viable oocysts of *Cryptosporidium parvum* in water after five freeze–thaw cycles. Robertson and Gjerde (2004) diluted a bovine fecal sample in water and found that initial numbers of *Giardia* cysts and *Cryptosporidium* oocysts were reduced by 0.6-log after five freeze–thaw cycles. Further research on the survival of *E. coli* and protozoa in winter conditions may therefore be valuable to investigate differences in ratios. The analysis of *E. coli* to protozoa ratios in rivers located in other climatic regions would also be relevant to investigate the nature and extent of this problem.

6. Conclusions

In this study, relationships between *E. coli, Cryptosporidium* and *Giardia* were quantified using paired *E. coli* and protozoa data collected in source water from 27 drinking water treatment plants (DWTPs). To evaluate the potential of *E. coli* for defining minimum treatment requirements, risk-based critical mean protozoa concentrations in source water were determined with a quantitative microbial risk assessment

(QMRA) model developed using the methods and assumptions recommended by the World Health Organization (WHO) Guidelines for drinking-water quality. This work led to the following conclusions:

- It was possible to evaluate site-specific mean *E. coli/Cryptosporidium* ratios and mean *E. coli/Giardia* ratios in source water by modeling the ratio of two correlated lognormal variables. Non-detects, correlations between *E. coli* and protozoa, and parametric uncertainties were taken into account with this model;
- *E. coli*/protozoa ratios at DWTPs supplied by small and medium rivers in agricultural or forested areas were typically 1.0 to 2.0-log lower than at DWTPs supplied by large rivers in urban areas. These results support the findings of Lalancette *et al.* (2014). The seasonal variation analysis revealed that these differences were related to low mean *E. coli* concentrations during winter at DWTPs supplied by small rivers;
- Power law regressions showed no proportionalities between the log of mean *E. coli* concentrations and the log of mean protozoa concentrations at the 27 sites. This analysis indicated that large changes in mean *E. coli* concentrations are related to small changes in mean *Cryptosporidium* or *Giardia* concentrations;
- Results from QMRA models indicated that a minimum reduction of 4.0-log would be needed at 20 DWTPs to achieve a health-based target of 10^{-6} DALY per person per year for *Cryptosporidium*. A minimum reduction of 4.0-log would be necessary at all DWTPs to achieve this target for *Giardia*;
- To achieve a target of 10^{-6} DALY per person per year, treatment requirements higher than 3.0-log would be needed at DWTPs with mean *E. coli* concentrations as low as 30 CFU 100 mL^{-1} for *Cryptosporidium* and 3 CFU 100 mL^{-1} for *Giardia*. Therefore, the definition of an *E. coli* trigger level would have limited value for defining health-based treatment requirements for protozoa at these DWTPs;
- Overall, this work suggests that *E. coli* monitoring does not provide valuable insight for the risk assessment of *Cryptosporidium* and *Giardia* at DWTPs supplied by small rivers in rural areas.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the NSERC Industrial Chair on Drinking Water, the Canadian Research Chair on Source Water Protection, NSERC Collaborative Research and Development Grant Project (CRDPJ-505651–16) and the Canada Foundation for Innovation. A part of the outcomes presented in this paper was based on research financed by the Dutch-Flemish Joint Research Programme for the Water Companies. We thank the technical staff of the biology and microbiology division at centre d'expertise en analyze environnementale du Québec and of the drinking water division at Ministère de l'Environnement et de la Lutte contre les changements climatiques for sample collection and laboratory analysis.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2021.117707.

References

Afolabi, E.O., Quilliam, R.S., Oliver, D.M., 2020. Impact of Freeze–Thaw cycles on die-off of E. coli and intestinal enterococci in deer and dairy faeces: implications for landscape contamination of watercourses. Int. J. Environ. Res. Public Health 17, 6999.

Burnet, J.-B., Sylvestre, É., Jalbert, J., Imbeault, S., Servais, P., Prévost, M., Dorner, S., 2019. Tracking the contribution of multiple raw and treated wastewater discharges at an urban drinking water supply using near real-time monitoring of β-d-glucuronidase activity. Water Res. 164, 114869.

Crow, E.L., Shimizu, K., 1987. Lognormal distributions. Marcel Dekker New York.

Garcia-Armisen, T., Servais, P., 2007. Respective contributions of point and non-point sources of *E. coli* and enterococci in a large urbanized watershed (the Seine river, France). J. Environ. Manage. 82, 512–518.

Gelman, A., Shirley, K., 2011. Inference from simulations and monitoring convergence. Handbook Markov Chain Monte Carlo 6, 163–174.

George, I., Anzil, A., Servais, P., 2004. Quantification of fecal coliform inputs to aquatic systems through soil leaching. Water Res. 38, 611–618.

Gibney, K.B., O'Toole, J., Sinclair, M., Leder, K., 2014. Disease burden of selected gastrointestinal pathogens in Australia, 2010. Int. J. Infect. Dis. 28, 176–185.

Gouvernement du Québec., 2012. Règlement et autres actes. Règlement modifiant le Règlement sur la qualité de l'eau potable (Loi sur la qualité de l'environnement). Gazette officielle du Québec. Pages 849-874Éditeur officiel du Québec 2012, editor22 févr. 2012.

Government of Alberta., 2012. Part 1: Standards for municipal waterworks of a total of 5 parts. Page 104 Standard and guidelines for municipal waterworks, wastewater and storm drainage systems. Edmonton, AB, Canada.

Government of Manitoba., 2007. The drinking water safety Act. Drinking water quality standards regulation (C.C.S.M. c D101). Manitoba, Canada. D101-M.R. 41/2007.

Haas, C.N., Rose, J.B., Gerba, C.P., 1999. Quantitative microbial risk assessment. John Wiley and Sons, Inc., New York, USA.

Havelaar, A.H., Melse, J.M., 2003. Quantifying public health risk in the WHO guidelines for drinking-water quality: a burden of disease approach. In: RIVM, 734301022. World Health Organization (WHO).

Health Canada. 2017. Enteric protozoa in drinking water: *Giardia* and *Cryptosporidium*. Federal-Provincial-TerritorialCommittee on Drinking Water.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1995. Continuous univariate distributions, volume 2. John wiley & sons.

Kass, R.E., Carlin, B.P., Gelman, A., Neal, R.M., 1998. Markov chain Monte Carlo in practice: a roundtable discussion. The American Statistician 52, 93–100.

Kato, S., Jenkins, M.B., Fogarty, E.A., Bowman, D.D., 2002. Effects of freeze–thaw events on the viability of Cryptosporidium parvum oocysts in soil. J. Parasitol. 88, 718–722.

Kruschke, J.K., 2014. Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan, 2nd edition. Amsterdam Academic Press. edition.

Lalancette, C., Généreux, M., Mailly, J., Servais, P., Côté, C., Michaud, A., Di Giovanni, G. D., Prévost, M., 2012. Total and infectious *Cryptosporidium oocyst* and total *Giardia cyst* concentrations from distinct agricultural and urban contamination sources in Eastern. Canada Journal of Water and Health 10, 147–160.

Lalancette, C., Papineau, I., Payment, P., Dorner, S., Servais, P., Barbeau, B., Di Giovanni, G., Prévost, M., 2014. Changes in *E. coli to Cryptosporidium* ratio from various fecal pollution sources and drinking water intakes. Water Res. 55, 150–161.

Lapen, D.R., Schmidt, P.J., Thomas, J.L., Edge, T.A., Flemming, C., Keithlin, J., Neumann, N., Pollari, F., Ruecker, N., Simhon, A., Topp, E., Wilkes, G., Pintar, K.D. M., 2016. Towards a more accurate quantitative assessment of seasonal *Cryptosporidium* infection risks in surface waters using species and genotype information. Water Res. 105, 625–637.

Loáiciga, H.A., Leipnik, R.B., 2005. Correlated gamma variables in the analysis of microbial densities in water. Adv. Water Res. 28, 329–335.

Madoux-Humery, A.-S., Dorner, S., Sauvé, S., Aboulfadl, K., Galarneau, M., Servais, P., Prévost, M., 2016. The effects of combined sewer overflow events on riverine sources of drinking water. Water Res. 92, 218–227.

Messner, M.J., Berger, P., 2016. *Cryptosporidium* infection risk: Results of new dose-response modeling. Risk Anal. 36, 1969–1982.

Nash, T.E., Herrington, D.A., Losonky, G.A., Levine, M.M., 1987. Experimental human infections with *Giardia lamblia*. J. Infect. Dis. 156, 974–984.

Ouattara, N.K., Passerat, J., Servais, P., 2011. Faecal contamination of water and sediment in the rivers of the Scheldt drainage network. Environ. Monit. Assess. 1–15.

Payment, P., Berte, A., Prévost, M., Ménard, B., Barbeau, B., 2000. Occurrence of pathogenic microorganisms in the Saint Lawrence River (Canada) and comparison of health risks for populations using it as their source of drinking water. Can. J. Microbiol. 46, 565–576.

Payment, P., Locas, A., 2011. Pathogens in water: value and limits of correlation with microbial indicators. Ground Water 49, 4–11.

Plummer, M. 2013. JAGS: Just Another Gibbs Sampler (Version 3.4. 0)[Computer software].

Regli, S., Rose, J.B., Haas, C.N., Gerba, C.P., 1991. Modeling the risk from Giardia and viruses in drinking water. J. Am. Water Works Assn. 83, 76–84.

Robertson, L., Gjerde, B., 2004. Effects of the Norwegian winter environment on Giardia cysts and Cryptosporidium oocysts. Microb. Ecol. 47, 359–365.

Schmidt, P.J., Chappell, C.L., 2016. *Cryptosporidium* infection risk: Results of new dose-response modeling" - Discussion of underlying assumptions and their implications. Risk Anal. 36, 2189–2192.

Smith, A.E., Ryan, P.B., Evans, J.S., 1992. The effect of neglecting correlations when propagating uncertainty and estimating the population distribution of risk. Risk Anal. 12, 467–474.

Swaffer, B., Abbott, H., King, B., van der Linden, L., Monis, P., 2018. Understanding human infectious *Cryptosporidium* risk in drinking water supply catchments. Water Res. 138, 282–292.

Sylvestre, E., Burnet, J.-B., Smeets, P., Medma, G., Prévost, M., Dorner, S., 2020a. Can routine monitoring of *E. coli* fully account for peak event concentrations at drinking water intakes in agricultural and urban rivers? Water Res. 170, 115369.

Sylvestre, É., Burnet, J.B., Dorner, S., Smeets, P., Medema, G., Villion, M., Hachad, M., Prévost, M., 2020b. Impact of Hydrometeorological events for the selection of parametric models for Protozoan Pathogens in drinking-water sources. Risk Anal.
Sylvestre, É., Prévost, M., Burnet, J.-B., Pang, X., Qiu, Y., Smeets, P., Medema, G., Hachad, M., Dorner, S., 2021. Demonstrating the reduction of enteric viruses by drinking water treatment during snowmelt episodes in urban areas. Water Research X.
Sylvestre, É., Prévost, M., Smeets, P., Medema, G., Burnet, J.B., Cantin, P., Villion, M., Robert, C., Dorner, S., 2020c. Importance of distributional forms for the assessment of Protozoan Pathogens concentrations in drinking-water sources. Risk Anal.
Teunis, P.F.M., Chappell, C.L., Okhuysen, P.C., 2002. Cryptosporidium dose-response studies: variation between hosts. Risk Anal. 22, 475–485.
Teunis, P.F.M., Rutjes, S.A., Westrell, T., de Roda Husman, A.M., 2009. Characterization of drinking water treatment for virus risk assessment. Water Res. 43, 395–404.
United States Environmental Protection Agency (USEPA), 2006. Federal Register. Part II. National primary drinking water regulations: Long Term 2 Enhanced Surface Water Treatment Rule; final rule. Office of Science and Technology. Office of Water, Washington, DC, USA.
USEPA, 2005a. Method 1623: Cryptosporidium and Giardia in water by filtration/IMS/FA. Office of Water, Washington, DC, USA. EPA 815-R-05-0024607.
USEPA, 2005b. Occurrence and Exposure Assessment For The Final Long Term 2 Enhanced Surface Water Treatment Rule. Office of Water. EPA 815-R-06-0024606-M.
USEPA, 2010. Long term 2 enhanced surface water treatment rule toolbox guidance manual. Office of Water, Washington, DC, USA. EPA 815-D-09-001.
USEPA. 2012. Method 1623.1: Cryptosporidium and Giardia in Water by Filtration/IMS/FA.
Wang, X., Zhang, D., Chen, W., Tao, J., Xu, M., Guo, P., 2019. Effects of fulvic acid and fulvic ions on Escherichia coli survival in river under repeated freeze-thaw cycles. Environ. Pollut. 247, 1100–1109.
WHO, 2009. Risk assessment of Cryptosporidium in drinking water. WHO/HSE/WSH/09.04, Public Health and Environment. Water, Sanitation, Hygiene and Health, Geneva, Switzerland.
WHO, 2016. Quantitative microbial risk assessment: Application for water safety management. Geneva, Switzerland.
WHO. 2017. Guidelines for drinking-water quality: Fourth edition incorporating the first addendum.
Wilkes, G., Ruecker, N.J., Neumann, N.F., Gannon, V.P.J., Jokinen, C., Sunohara, M., Topp, E., Pintar, K.D.M., Edge, T.A., Lapen, D.R., 2013. Spatiotemporal analysis of cryptosporidium species/genotypes and relationships with other zoonotic pathogens in surface water from mixed-use watersheds. Appl. Environ. Microbiol. 79, 434–448.
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. J. Water Health 9, 265–278.