Environmental Science & Technology

Policy Analysis

pubs.acs.org/est

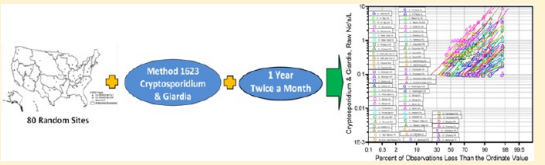
ICR SS Protozoan Data Site-by-Site: A Picture of *Cryptosporidium* and *Giardia* in U.S. Surface Water

Jerry E. Ongerth*

Civil, Mining, & Environmental Engineering, University of Wollongong, Wollongong, New South Wales 2522, Australia

Supporting Information

ABSTRACT: The U.S. Environmental Protection Agency (USEPA) Information Collection Rule Supplemental Survey (ICR SS) required analysis of *Cryptosporidium* and *Giardia* in 10 L surface water samples twice a week for a year by USEPA Method 1623 at 80 representative U.S. public water systems (PWS). The resulting data are examined site-by-site in relation to objectives of the Federal drinking water regulation, The Long-Term (2) Enhanced Surface Water Treatment Rule (LT2), currently under formal 6-year review by the USEPA. The data describe *Cryptosporidium* and *Giardia* in watersheds nationwide over a single annual cycle. Due to limited recovery efficiency measurement results are not fully quantified. In the required sample volumes of 10 L no *Cryptosporidium* were found in 86% of samples and no *Giardia* were found in 67% of samples. Yet, organisms were found in enough samples at 34 of 80 sites to detail a specrtum of occurrence and variability for both organisms. The data are shown to describe indivudual site risk essential for guidance of watershed and water treatment management by PWSs. The span of median occurrence for both organisms was about 2 orders of magnitude above the limit of detection (LD), ca. 0.05 raw no's/L for *Cryptosporidium* and ca. 0.10 raw no's/L for *Giardia*. Data analysis illustrates key features of *Cryptosporidium* and *Giardia* in surface water: presence is continuous not intermittent; zeros indicate presence below the LD; occurrence level and variations depend on watershed sources; risk depends on both magnitude and variability of concentration; accurate estimation of risk requires routine measurement of recovery efficiency and calculation of concentration. The data and analysis illustrate features of *Cryptosporidium* and *Giardia* occurrence in surface water relevant to their effective regulation for public health protection.



1. INTRODUCTION

The waterborne protozoan pathogens *Cryptosporidium* and *Giardia* are the subject of Federal drinking water regulation in the U.S., the Long-term 2 Enhanced Surface Water Treatment Rule (LT2),[1] including formal Federal requirement for data collection, the Information Collection Rule (ICR).[2] Data on *Cryptosporidium* and *Giardia* occurrence in representative surface water sources used by public water systems (PWS's) across the U.S. were produced in an addition to the Information Collection Rule called the Supplemental Survey (ICR SS).[2] Site-by-site analysis of this unique and extensive data set is used to raise timely issues relevant to drinking water regulation in the U.S. Significant work was done by the U.S. Environmental Protection Agency (USEPA) and contractors using data from the ICR and the ICR SS in developing and finalizing provisions of LT2.[1] However, in generalizing the negative-result-dominated data for the purpose of formulating monitoring requirements for LT2, implications of the data for each PWS and their individual sampling sites were not thoroughly explored.[3] The focus of the analysis in this paper is the ICR SS individual site data and their implications. In reviewing these data along with recently published analysis of the LT2 data some fundamental issues regarding the formulation of the Surface Water Treatment Rule are raised.

The critical issue addressed by LT2 and to which the *Cryptosporidium* (and *Giardia*) monitoring of both the ICR SS and LT2 were directed is to identify the relative risk associated with these organisms to surface water-using PWS's.[4] The rule includes risk management in four categories or "BINs" based on LT2 monitoring data,[3] imposing incremental *Cryptosporidium* control requirements for increasingly higher BINs. Recognizing the principle that risk is proportional to concentration, previous work on these organisms has established that the concentration of *Cryptosporidium* and *Giardia* is generally proportional to the extent and intensity of organism generating activity in the tributary watershed.[5,6] Hence, concentrations and risk should logically increase downstream with increasing watershed area and activities including domestic animal raising and human population. An overriding shortcoming of the ICR SS data is that recovery efficiencies were not measured nor applied to calculate concentrations from the raw numbers of organisms found.

The concentrations of *Cryptosporidium* and *Giardia* in surface water in the U.S. are of particular importance in 2013. This is due to the LT2 requirement for a second round of *Cryptosporidium* monitoring by the same PWS's beginning in 2016.[1] In addition the LT2 regulation is in the midst of a

Received: June 24, 2013
Revised: August 13, 2013
Accepted: August 14, 2013

   dx.doi.org/10.1021/es4027503 | Environ. Sci. Technol. XXXX, XXX, XXX–XXX

Federally required formal Regulatory Review process, concurrent although coincidental with timing of the LT2 second phase of monitoring, (see Supporting Information (SI) Figure A [7]). The first round of LT2 monitoring produced an extensive body of data on *Cryptosporidium* occurrence at surface water intake locations of public water systems (PWS's) serving populations >10 000 in the period from 2006 to 2010.[4] *Giardia* analysis was not included in LT2 monitoring. Results of the first round of LT2 monitoring were summarized by the EPA[8] and released for public use in mid 2012, http://water.epa.gov/lawsregs/rulesregs/sdwa/lt2/upload/cryptodatareported.csv. The EPA summary (see SI Figure B), described the results of about 45 000 samples from nearly 1700 sample sites in 1250 PWS's across the U.S. Included were the results of nearly 40 000 field samples and over 3000 matrix spiked samples that were required for quality control. More than half of the PWS's reported finding no *Cryptosporidium* in any of the required 24 consecutive monthly samples. Nation-wide, in 93% of all field samples analyzed no *Cryptosporidium* were found. Yet, at over 100 sampling sites in every region of the country, *Cryptosporidium* were found in enough samples to permit estimating the principal occurrence characteristics defining risk, that is, the median concentration and degree of variability, Figure 1.[9,10] When analyzed on a site-by-site basis, the LT2 data

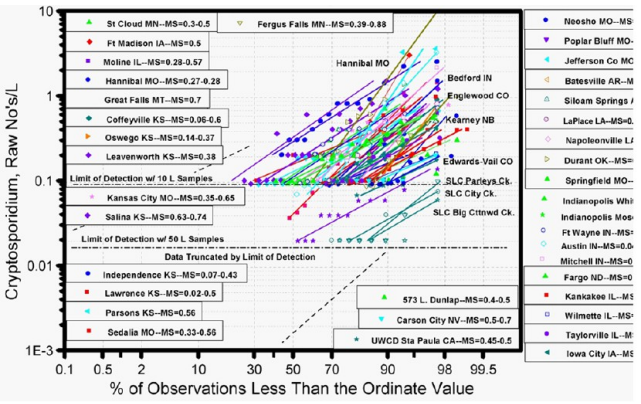


Figure 1. Cumulative frequency of LT2 *Cryptosporidium* occurrence at 50 representative sampling locations in regions across the U.S.[9].

describe a spectrum of occurrence from a minimum in western mountain streams to a maximum in more heavily populated areas along the larger river systems in the Midwest and eastern parts of the country. The degree of variability in *Cryptosporidium* occurrence is shown by the slope of the data from each sampling site and had a relatively narrow but significant range. Variability appeared least in downstream locations on large river systems and greatest in upstream locations on smaller streams.

In light of this background, data produced by the Supplemental Survey of the Information Collection Rule[2] collected between March 1999 and February 2000, were examined. The ICR SS data file was provided by EPA on request in a Google Docs file (db2.mdb) at https://docs.google.com/open?id=0B7tGKe6V6-BeNTRBN3dFMlFPTFE.

2.0. MATERIALS AND METHODS

Methods and procedures for this work included elements of the Information Collection Rule that specified requirements for data to be collected and analytical procedure affecting the resulting data.[2] Also described are procedures used here in analyzing the ICR SS data.

2.1. USEPA ICR Supplemental Survey data. The ICR SS was conducted to take advantage of the significantly improved analytical method for *Cryptosporidium* and *Giardia*, USEPA Method 1622 and 1623,[11] developed and validated during the ICR period. The ICR SS was conducted at 87 sites, 7 "certainty sites" among the largest water treatment plants in the U.S., and 80 randomly selected PWS using surface water at locations throughout the U.S. 40 "large" plants serving populations >100 000, and 40 "medium" plants serving populations between 10 000 and 100 000, Figure 2. Plants in both large and medium

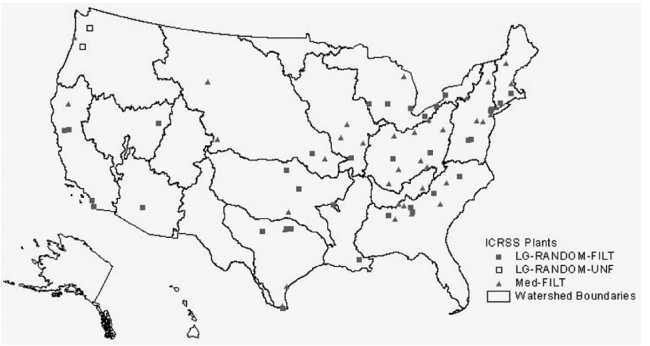


Figure 2. Location of ICR SS sampling sites in the U.S. with major hydrologic region boundaries.[3].

categories were selected to represent equally flowing streams and reservoir/lake sources. The ICR SS data collection was designed to meet four objectives:[3] primarily, (1) to characterize (a) the national distribution of *Cryptosporidium* and *Giardia* occurrence and (b) the features of occurrence at individual plants (mean, median, and 90th percentile); (2) to compare the national distributions between large and medium systems; and secondarily, (1) to identify differences in occurrence between flowing stream and reservoir/lake sources and the potential association with watershed features characterized by coliform concentration; and (2) to examine occurrence in relation to other ancillary water quality parameters.

Participating PWSs were required to monitor biweekly for 12 months. For the first 4 months 10 L samples were analyzed by Method 1622 for *Cryptosporidium*. For the remaining 8 months the 10 L samples were analyzed by Method 1623 for both *Cryptosporidium* and *Giardia*. The analytical methods result in reporting microscopy observations that divide the total oocyst and cyst numbers identified into categories based on evidence of internal contents potentially associated with viability. The resulting data were compiled by EPA in the form of a Microsoft Access Database, noted above. The file of 5.4 MB includes a dozen tables containing the primary protozoa data, data on coliform concentrations, ancillary water quality parameter data, PWS location data, and explanatory information.

2.2. USEPA Analysis of the ICR SS Data. The USEPA conducted extensive analysis of Supplemental Survey data from the 80 randomly selected sites, directed to satisfying their objectives summarized above and detailed in Chapters 3 and 4 of the Occurrence and Exposure Assessment.[3] A major focus of the analysis was modeling of the observed data to estimate what would be observed in data collection to be required as part of the LT2 regulation, then under development.

dx.doi.org/10.1021/es4027503 | Environ. Sci. Technol. XXXX, XXX, XXX−XXX

2.3. Analysis of the USEPA ICR SS data file. Analysis of the USEPA ICR SS data for this paper was performed using Microsoft Excel functions to sort the file to generate sub files containing the data for individual sampling points at which four or more field samples contained *Cryptosporidium* oocysts and/or *Giardia* cysts. All subsequent analysis used only the total numbers of both organisms as indicative of the source and hence risk potential at each individual sampling point. Microsoft Excel tools were used for simple plotting and calculations, for example, standard deviation (SD) and coefficients of variation (cv) where cv = SD ÷ mean).

The sorted protozoa data were scanned manually to identify sampling sites at which positive findings of *Cryptosporidium* and *Giardia* were sufficient to permit further analysis by forming log-probability distributions as previously described.[12] The resulting distributions provide the key features related to risk, that is, median concentration and degree of variability.[5,6] *Cryptosporidium* and *Giardia* occurrence distributions for individual sampling sites were formed using OriginPro 8.6 (OriginLab, Northhampton, MA) as previously described.[12] Briefly, numbers of oocysts reported were normalized by sample volume and ranked high to low. For individual sites the ranked occurrence/L Excel tables including zeros were exported to OriginPro 8.6 to generate log probability plots as previously described using the "probability plot" function under plot → statistics.[12]

3.0. ICR SS DATA AND ANALYSIS

3.1. USEPA ICR SS Data and Analysis Summary. In the EPA Occurrence and Exposure Assessment,[3] the Supplemental Survey data were analyzed in accordance with the survey objectives. Statistical analysis of the total data set and individual site data was conducted principally "...to characterize the occurrence of *Cryptosporidium* and other pathogens in surface water used..." for public water supply. A major facet of the analysis was modeling the data to permit predicting the nationwide occurrence distribution for assignment of risk-proportional BINs.[3] In a December 2011 USEPA LT2 Stakeholder Meeting presentation[8] the LT2 data were summarized in relation to the Supplemental Survey data for *Cryptosporidium* indicating significant overestimation of predicted vs observed occurrence in LT2 monitoring results (see SI Figure B):

- Only 4.7% of PWS sites had occurrence greater than the lowest BIN level compared to the 14% prediction:
- Only 7% of 40 000 LT2 source water measurements found *Cryptosporidium* compared to 14% of 2086 in the ICR SS, that is, 93% LT2 field samples were zeros;
- The average LT2 *Cryptosporidium* occurrence was 0.016/L compared to 0.053/L in the ICR SS. (Both computed by dividing the total number of *Cryptosporidium* by the total volume analyzed);
- Over 51% of 1760 sites in LT2 reported all zeros vs 23% of the 80 ICR SS plants.

3.1. ICR SS Data and Analysis Summary. Combining all 87 sampling locations in the ICR SS, a total of 2086 samples were analyzed for *Cryptosporidium* in which a total of 1057 oocysts were found in 283 of the 2086 (13%) samples. No *Cryptosporidium* were found in any of the samples from 18 of the 87 (20%) plants surveyed. A total of 1350 samples were analyzed for *Giardia* in which a total of 3255 cysts were found in 445 of 1350 (33%) samples. Completely negative findings for *Giardia* occurred in 16 of the 87 (18%) plants surveyed. In the 12 months of ICR SS sampling a total of 432 MS samples, were analyzed for *Cryptosporidium*, 5 per PWS, and 269 MS samples were analyzed for *Giardia*, 3 per PWS. The sample volumes analyzed were virtually all 10 L ± 1 except for smaller volumes limited by high turbidity.

The aggregate of all positive findings by month, Figure 3, shows that *Cryptosporidium* occurrence was relatively high in

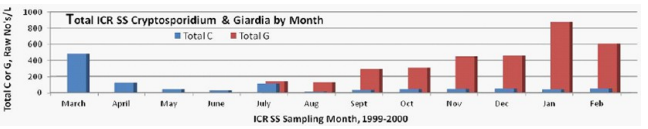


Figure 3. ICR SS *Cryptosporidium* and *Giardia* monthly totals, March 1999 to February 2000.

the spring months, March and April but virtually constant for the rest of the annual cycle. *Giardia* increased through the end of the year.

Each PWS was asked to analyze five matrix spike (MS) samples for *Cryptosporidium* and three for *Giardia*. The measurements were used only as a quality control indicator. Variations in MS recovery efficiency were characteristically high[10] both for individual sites and overall, Figure 4a and b. Overall including all 432 MS's for *Cryptosporidium* and 269 MS's for *Giardia*, recovery efficiencies for both *Cryptosporidium* and *Giardia* ranged from virtually nil to over 100%. The average MS recovery for *Cryptosporidium* was 43% with SD = 21% and cv = 48%. For *Giardia* the average MS recovery was 53% with SD = 24% and cv = 45%. Modest trends can be seen by eye, increasing in early spring for *Cryptosporidium* with generally lower values through summer and autumn, and for *Giardia* increasing through summer into winter. However, data for individual sites are insufficient to support statistical analysis to identify seasonality. Combining measurements from unrelated sites for statistical analysis would be inappropriate.

Of the 80 large and medium PWS surface water sources surveyed 34 reported 4 or more positive findings for either *Cryptosporidium* or *Giardia* and in 28 of 34 sites, both, Figure 5. The distributions of both *Cryptosporidium* oocyst and *Giardia* cyst occurrence were approximately log-normal. All distributions were truncated by the limit of detection imposed by the approximately 10L sample volumes analyzed. Median levels for *Cryptosporidium* fell below the limit of detection (LD) for all but 2 of the 34 sites. Median levels for *Giardia* were below the LD for about half of the 34 sites. By extrapolation median occurrence (Occ.) levels of both organisms covered a range of about 2-logs from about 0.002 to 0.2 raw no's/L (Table 1). Slopes of individual distributions indicating the degree of variability of occurrence for both organisms fell within a relatively narrow range of about 1 order of magnitude as indicated by relative standard deviations (RSD's) 1 < RSD < 10, where

$$\mathrm{RSD} = ([\mathrm{Occ.}]_{86.16\%\mathrm{ile}} - [\mathrm{Occ.}]_{50\%\mathrm{ile}})/[\mathrm{Occ.}]_{50\%\mathrm{ile}} \quad (1)$$

For *Cryptosporidium*, except for two higher values, of 115 and 22 (Reading PA and N. Bay CA respectively), the average RSD was about 4 with an equal SD (Table 1). Resulting mean occurrences for *Cryptosporidium* calculated from the extrapolated medians and graphically estimated standard deviations (SD's), eq 2:[13]

dx.doi.org/10.1021/es4027503 | Environ. Sci. Technol. XXXX, XXX, XXX–XXX

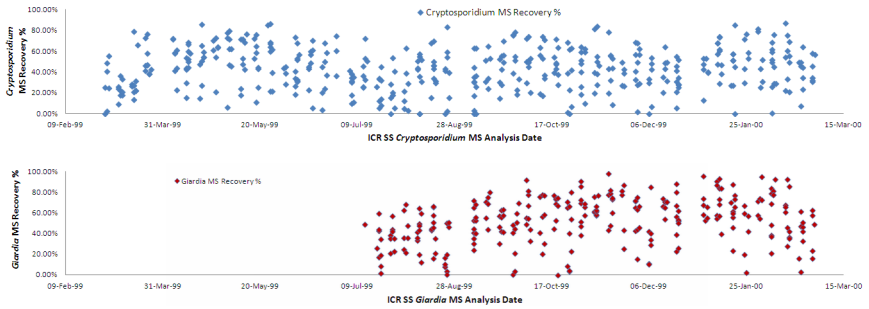


Figure 4. ICR SS matrix spike recovery efficiency % listed chronologically by sampling location for *Cryptosporidium* (a, blue) and *Giardia* (b, red).

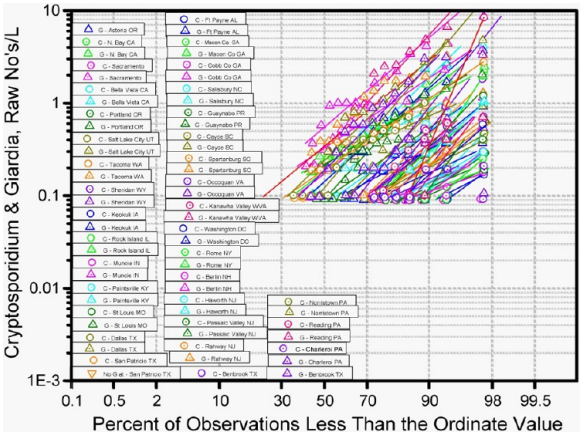


Figure 5. Cumulative frequency distributions of ICR SS *Cryptosporidium* and *Giardia* occurrence, raw no's/L, for 34 PWS sites surveyed reporting 4 or more nonzero results.

$$m_e = m_d \exp(2.7509(SD_{\log_{10}})^2) \quad (2)$$

where $SD_{\log_{10}}$ is the standard deviation of the log distribution determined graphically as eq 1:

$$SD = [\text{occ.}]_{(50+34.16)\%\text{ile}} - [\text{occ.}]_{50\%\text{ile}} \quad (3)$$

ranged from 0.03/L to 70/L but much more narrowly excepting the two highest RSD values. The *Giardia* occurrence ranged from 0.004 to 0.56/L with an average of 0.13/L and SD of about 0.13 and cv of nearly 100% (Table 1). The variability in *Giardia* occurrence at individual sites indicated by the RSD averaged about 6.3 with an SD of 7.8 and cv of about 125%. The calculated mean *Giardia* occurrence ranged from 0.06 to 7.8/L with an average of about 1/L, SD = 1.6, and cv of 160% (Table 1).

Occurrence distributions grouped for *Cryptosporidium* and *Giardia* individually show the generally higher median occurrence of *Giardia* (Figure 6b) than *Cryptosporidium* (Figure 6a), but a somewhat greater range of variability for *Cryptosporidium* than for *Giardia*.

The relative risk was ranked for each of the 34 sites for both the mean occurrence and degree of variability (RSD). An aggregate risk was determined as the sum of mean and RSD ranks (mean occ. rank + RSD rank). The resulting risk order for *Cryptosporidium*, Figure 7a includes sites largely in the mid Atlantic and Midwest at the high end and predominantly mountain west sites near the low end, although with many exceptions. The risk order suggested by this ranking for *Giardia*, Figure 7b, is more mixed in terms of site locations with the highest half of sites located in the East, South and Central regions of the country. Grouping the occurrence distributions separately for *Cryptosporidium* and *Giardia* into four regions, Central, East, South, and West, Figure 8a–h, the apparent differences in occurrence are shown more clearly. For *Cryptosporidium*, of the 34 sites only two in the Eastern grouping had median occurrence > LD, 0.1 Raw no's/L. For *Giardia*, 3 of 8 Central sites, 7 of 9 Eastern sites, and 4 of 8 Southern sites had median occurrence > LD.

4.0. DISCUSSION

Throughout the literature data on *Cryptosporidium* and on *Giardia* in water with few exceptions rarely include more than a few measurements at a single site.[14−16] This is partly due to the time, effort, and expense required in sampling and analysis, and partly due to the nature of projects that generated the data. Thus, data produced in the ICR SS (87 sites × 24 *Cryptosporidium* and 16 *Giardia* analyses, in a single 12 month period) were unparalleled in their contribution to understanding the occurrence of these organisms in water. The challenge is to understand just what the data tell us in order to direct the most effective water quality control for PWS's and the most effective regulatory guidance. The most important questions affecting individual PWS's, include ones related to

Table 1. ICR SS *Cryptosporidium* and *Giardia* Occurrence Characteristics Estimated from Cumulative Probability Distributions of 34 Sites Having Four or More Non-Zero Organism Values

Cryptosporidium	m_d, no/L	RSD	m_e,No/L	*Giardia*	m_d, No/L	RSD	m_e, no/L
max	0.22	115	70.17	max	0.56	42.02	7.79
min	0.00086	0	0.03	min	0.01	0	0.06
avg	0.048	7.81	2.31	avg	0.13	6.34	1.03
SD	0.050	19.65	12	SD	0.13	7.83	1.64
cv	1	2.52	5.20	cv	0.98	1.23	1.59
avg w/o 2 highest		3.98	0.16				
SD w/o 2 highest		4.00	0.17				
cv w/o 2 highest		1.01	1.09				

dx.doi.org/10.1021/es4027503 | *Environ. Sci. Technol.* XXXX, XXX, XXX−XXX

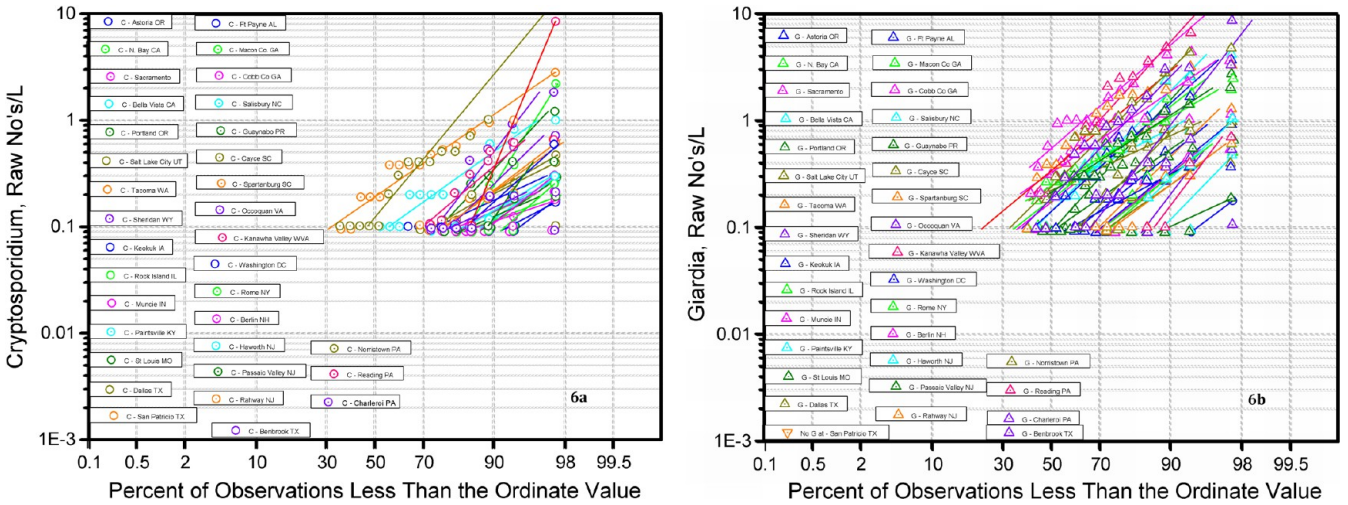


Figure 6. Cumulative frequency distributions of ICR SS *Cryptosporidium* occurrence (6a) and *Giardia* occurrence (6b), both raw no's/L.

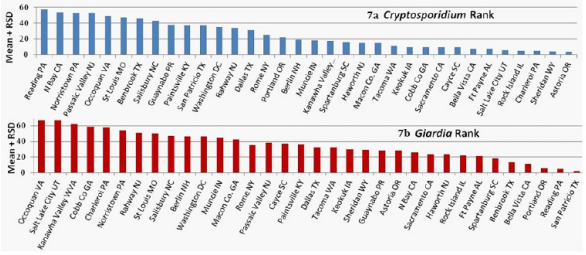


Figure 7. Relative risk ranking by mean occurrence and RSD of *Cryptosporidium* and *Giardia* for 34 ICR SS sites having four or more nonzero sample results.

relative risk, watershed management, and management of water treatment to meet water quality and public health requirements:

1. At what level (expressed as concentration taking analytical recovery into account) do *Cryptosporidium* and *Giardia* occur in the water at any specific sampling point?
2. How do levels of *Cryptosporidium* and *Giardia* at individual sampling points vary over typical annual cycles?
3. To what degree does recovery efficiency at individual sampling points vary over typical annual cycles and how does its variation relate to variations in organism occurrence?
4. What are the likely sources of the organisms in the watershed and how are they likely to affect the level and variations in organism concentrations at the point of diversion?
5. Are the data observed at any specific sampling point consistent with data from sampling sites having similar watershed and organism source characteristics elsewhere?

Public policy issues pertaining to development and implementation of a regulatory system to assist PWS's in control of *Cryptosporidium* and *Giardia* all rest on the fundamental principle that public health risk due to these organisms alone is directly proportional to their concentration at the point of consumption. That concentration is driven directly, through the complex sequence of treatment including physical removal and inactivation processes, by the concentration at the point from which the surface water is diverted into the PWS. Accordingly, measurement of concentration is the fundamental requirement. Also, it is argued here that total concentration, rather than ephemeral subsets of "viable" and or "infectious" organisms, is the most important information both in terms of risk potential and the PWS's ability to manage both source water quality and treatment effectiveness.

Before discussing the ICR SS data relevant to the questions above it is essential that the reader understand and appreciate that the data collected in the ICR SS are not measurements of concentration. Due to the additional effort and expense required to produce concentration data by measurement of recovery efficiency relevant to every sample analyzed with few exceptions (e.g., refs 5, 6, and 10) virtually all data in the literature on *Cryptosporidium* and *Giardia* in water, including those of the ICR SS, are limited to raw numbers/L. The typically systematic variation of recovery efficiency, independent of variations in organism occurrence, and with variations due to annual water quality cycles unique to virtually every surface water sampling site makes interpretation of data on these organisms without recovery efficiency difficult and often misleading.[12] The relatively high variation in MS measurements for each the 87 individual ICR SS sampling sites is indicated by the coefficients of variation (cv = the standard deviation ÷ the mean) of the five MS measurements for each site, Figure 9. The site by site cv's ranged from a minimum of 10% to over 120%, averaging 45% for *Cryptosporidium*, and averaging 41% for *Giardia* with a range from 5% to 137%. This degree of variation in the recovery efficiency for *Cryptosporidium* is consistent with previously reported observations.[12,17] The variations in MS recoveries at individual sites as observed in the ICR SS data raise the question of reproducibility. Recently reported triplicate measurements of MS recovery for both *Cryptosporidium* and *Giardia* by Method 1623 indicate that they are reproducible.[12] Thus, observed variability in MS measurements at individual sites should be interpreted strictly as a systematic matrix effect varying at individual sites over typical annual water quality cycles at each site, precisely what the MS tool in Method 1623 is designed to describe and take into account.

Although the ICR SS data cannot be expressed in terms of concentration, the occurrence data provide indications of the overall distribution of both *Cryptosporidium* and *Giardia* across the U.S. These data provide the earliest detailed picture of

dx.doi.org/10.1021/es4027503 | *Environ. Sci. Technol.* XXXX, XXX, XXX–XXX

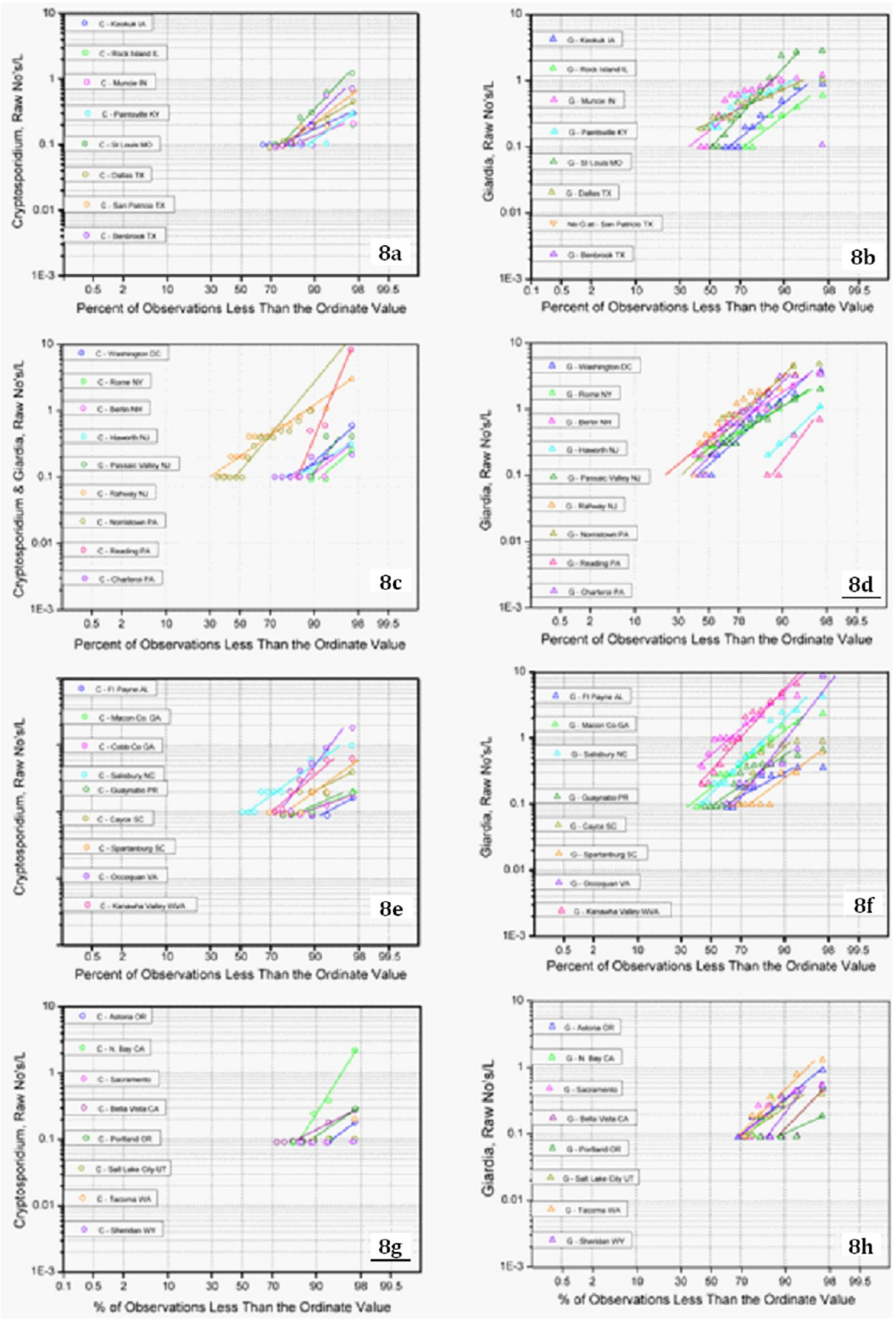


Figure 8. Cumulative frequency distributions of ICR SS *Cryptosporidium* (a, c, e, g) and *Giardia* (b, d, f, h) occurrence, raw no's/L, for regional site groupings; Central (a, b), East (c, d), South (e, f), and West (g, h).

occurrence at a broad range of individual sampling points during the same sampling period. As observed above no data reported previously are comparable. As summarized by the USEPA, no *Cryptosporidium* were detected in 86% of samples analyzed and no *Giardia* were detected in 67% of samples analyzed.[18] The average observed occurrences for *Cryptosporidium* and *Giardia* including all zeros were ca. 0.05 per L and 0.35 per L, respectively, Figure 10. However, this is of little relevance to an individual PWS and overlooks the information in the positive findings at nearly half (34 of 80) of the sites sampled, Figure 6a and b, and Table 1. As shown by these figures and in more detail on Figure 8a–h, *Cryptosporidium* and

dx.doi.org/10.1021/es4027503 | *Environ. Sci. Technol.* XXXX, XXX, XXX–XXX

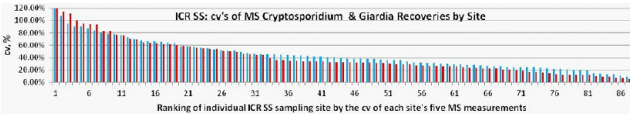


Figure 9. Ranking of the cv's of the MS measurements (five for *Cryptosporidium*, three for *Giardia*) per site for the 87 ICR SS sampling sites.

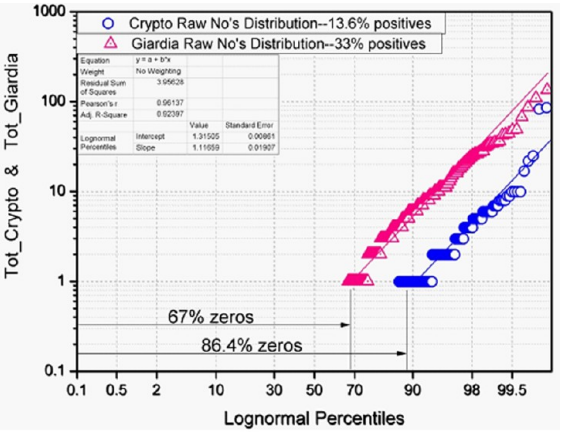


Figure 10. Cumulative frequency distributions of all positive ICR SS sample analyses by Method 1622 and 1623 for *Cryptosporidium* and *Giardia*, raw nos./L.

Giardia occurrence at individual sites have characteristic median levels and degrees of variability unique to the combination of watershed and surface water quality characteristics at each individual sampling site. Limited comparisons can be made between sites although much caution must be used due to the lack of recovery measurements and the inability to account for their variation independent of variations in organism occurrence. The combined data from all 34 sites, for *Cryptosporidium* occurrence, Figure 6a, and for *Giardia* occurrence, Figure 6b, suggest a spectrum of occurrence in surface waters ranging in quality from high mountain streams to downstream locations on the country's major river systems. Anomalies relative to intuitive assessment are apparent in risk ranking according to median occurrence and variability for both *Cryptosporidium* and *Giardia*, Figure 7a and b. Including recovery measurement and use of it to compute concentrations would likely resolve such anomalies. The true spectrum that would result if recovery was measured consistently producing concentration distributions, would satisfy the intended objective of the ICR and LT2 monitoring, that is, to identify the true spectrum of *Cryptosporidium*-related risk at each surface water intake location of PWS across the country.

Comparison of occurrence data between available data sets may be useful. Some significant data on the occurrence (raw nos./L) of both *Cryptosporidium* and *Giardia* were collected during the same 1999–2000 period of the ICR SS but independently.[15] The analytical method used was the previous ICR method reported most often as having lower and more variable recovery than Method 1622/1623.[3] Nevertheless, in 100 L samples collected from 66 water treatment plants (not identified other than by State) located predominantly in the Central and Northeast U.S., both *Cryptosporidium* and *Giardia* were found widely distributed. *Cryptosporidium* were found in 74 of 85 (87%) of samples at an average occurrence of 7.3 raw no's/L, range 0.04 to 66, SD = 10, and *Giardia* were found in 69 of 85 (81%) of samples at an average occurrence of 7.5 raw no's/L, range 0.04 to 74, SD = 14 (neglecting 3 values >100). The higher occurrence at these sites similar to about half of the ICR SS sites cannot be explained. Occurrence at these sites averaged more than 10× that of the ICR SS data, Figure 11, in

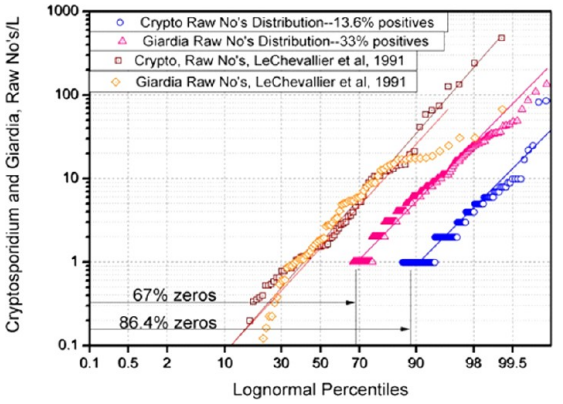


Figure 11. Cumulative frequency distribution of *Cryptosporidium* and *Giardia* occurrence previously reported[15] in comparison to ICR SS data distributions (Figure 10).

spite of having resulted from a generally less efficient analytical method. The occurrence of *Cryptosporidium* and *Giardia* in similar ranges, where in other data *Giardia* have typically been found at higher levels than *Cryptosporidium*, Figure 10, is also anomalous. The reported occurrences at levels below 1/L are a result of having analyzed 100 L sample volumes providing a nominally 10-fold lower limit of detection. These data further reinforce the interpretation of zero findings as simply below the limit of detection, and that the presence of both organisms in surface water is continuous and not intermittent.

An additional comparison to previous reports can be made for the ICR SS sites of Rome NY and Berlin NH. Both are sites among the earliest of waterborne giardiasis outbreaks reported in the literature.[19,20] The ICR SS data from those locations show that *Giardia* occurrence was relatively high, averaging 0.26 and 0.33 raw no's/L, respectively, Figure 8d, each also with appreciable *Cryptosporidium* occurrence, Figure 8c.

The ICR SS data as described above provide some direction toward understanding occurrence of *Cryptosporidium* and *Giardia* in surface water worldwide. It is clear that both organisms occur in surface water at levels detectable using EPA Method 1622/1623. Full understanding of both the levels and variations over a typical annual cycle will not be possible without measuring recovery efficiency and using the measurements to calculate organism concentrations. The range of occurrence for most sites appears to be ca. 2-logs with higher levels more likely downstream of more highly developed watershed areas.

The last two major questions of interest in the data identified above deal with the relation between observed data and independent knowledge of likely sources of the organisms and processes occurring between the watershed and sampling points that would account for observed occurrence and variability. Such questions require additional information not available here. The information sources do exist in sanitary surveys (required of PWS's under the IESWTR) along with water quality data and limited estimates of *Cryptosporidium* and *Giardia* source production rates per unit area and time published previously.[5,6,10] These literature references used

dx.doi.org/10.1021/es4027503 | *Environ. Sci. Technol.* XXXX, XXX, XXX–XXX

Giardia and *Cryptosporidium* concentration data, having measured recovery efficiency and used it to calculate concentrations from observed numbers/L, along with watershed areas and streamflow gauging data to estimate *Giardia* and *Cryptosporidium* production from undeveloped, controlled, public water supply watersheds on the order of 10^6–10^7 organisms per mi^2/day. These production rates from controlled watersheds having only wild animal sources may serve as a useful baseline. Progressively higher rates would be expected to result from watersheds having progressively higher levels of development and sources including domestic animal production.

It is not clear why no organisms, neither *Cryptosporidium* nor *Giardia* nor both, were found at a significant proportion of sites.[18] Source characteristics of sites at which no organisms were found do not appear to differ significantly from those of sites at which positive findings occurred in a significant proportion of samples. It is also not clear why significantly fewer positive findings occurred and at lower average levels in LT2 sampling compared to that of the ICR SS. Factors likely contributing to answering these questions include: 1) minimal limit of detection due to minimal sample volumes; 2) lack of relevant recovery efficiency measurement; 3) large-scale programs taxing analytical resources and essential feedback of analytical results to individual sampling programs; and 4) disincentive related to maximum contaminant level goals (MCLG's) of zero and the prospect of large capital expenditure required of PWS having higher than BIN 1 occurrence.

The mandatory 6-year review mechanism imposed in the 1996 Safe Drinking Water Act Amendments[21] provides an opportunity for regulatory staff and water industry representatives to digest the compiled information provided by the ICR SS and LT2. The principal element that appears not to have been clearly identified so far though the monitoring and BIN calculation process is the accurate identification of surface water source locations in risk categories truly proportional to *Cryptosporidium* (and *Giardia*) presence. The spectrum of occurrence identified by the individual site data of the ICR SS (Figure 5) and the similar spectrum of occurrence in the LT2 data (Figure 1) provide clear evidence of virtually universal distribution of these organisms at levels that are readily detectable.[10] It must be acknowledged that if the monitoring regime were altered to remove disincentive for finding organisms, larger samples were analyzed, and recovery efficiencies measured to permit defining concentration, the real occurrence and hence exposure of the population would not be changed...only the frame of reference. Nevertheless, if the true spectrum of risk due to these organisms is to be established enabling individual PWSs to understand their true position within the risk spectrum it can be accomplished as described above. Allocation of additional funds to monitoring for *Cryptosporidium* and *Giardia* will not produce useful information if pursued as previous programs yielding mostly negative results. Defining true concentrations and their variation over typical annual cycles would permit comparison of data from one location to another and from one annual cycle to any other. Logical attention to available risk management tools can follow.

In terms of public policy the most interesting feature of the ICR SS data is the apparent description of a spectrum of occurrence for both *Cryptosporidium* and *Giardia*. The principal objective of the LT2 regulation is to identify those PWSs at higher risk and to apply corresponding risk management. The ICR SS data, Figure 5, derived from 34 of 87 PWS representative of surface waters across the country appear to describe the range of *Cryptosporidium* and *Giardia* occurrence. Neglecting the lack of recovery efficiency-derived concentration data, and without taking the degree of variability into account the occurrence range should also describe the range of relative risk. A major question of this logic is the 53 (87–34) PWS at which *Cryptosporidium* and *Giardia* were not found. Two possible answers include (1) No *Cryptosporidium* and *Giardia* were present at the all zero sites; or (2) The *Cryptosporidium* and *Giardia* occurrence at those sites was lower, that is, an extension of the spectrum described in Figure 5 (or Figure 6a and b). The occurrence of *Cryptosporidium* and/or *Giardia* at lower levels at sites having analyzed larger samples along with previous reports specifically testing the first question argues emphatically that zeros do not mean absence...simply occurrence below the limit of detection. The data from LT2 sampling sites for which 50 L samples were analyzed, shown in Figure 1, provide direct examples of this. If the data from 50 L samples, predominantly from relatively undeveloped watersheds, are taken as representing the lower extremity of occurrence, the possibility of sites having even lower levels of occurrence, particularly downstream of watersheds other than pristine, seems remote. This should be particularly true for PWSs on surface waters of generally lower quality and derived from watersheds having more extensive sources of *Cryptosporidium* and *Giardia*.

As the USEPA proceeds with the 6-year review of LT2, the individual site data from both ICR SS and LT2 phase 1 as described here provide a perspective not included in previous analysis. Deciphering the anomalies and resolution of clearly answerable questions raised by the data can lead to productive adjustments in the regulation. The desirable result would be more accurate and defendable public health protection. Under the existing rule, the prospect of significant capital cost associated with finding *Cryptosporidium* above the LT2 BIN 1 level combined with the unrealistic and technically undefendable MCLGs of zero combine to exert strong disincentive for any PWS to collect data other than zeros. The extremely high proportion of negative (zero) analytical results from ICR SS and LT2 appears to be at least in part if not mostly a reflection of this. Without suggesting collusive activity, any technician experienced with sampling and analysis for *Cryptosporidium* and *Giardia* understands the challenging nature of the needle-in-a-haystack process and the effort required to find organisms that are present or conversely, the ease of finding nothing. Finding zeros has been so common...and apparently acceptable...that simple checking procedures such as intersite comparisons, comparisons to previous data from the same site, adjustment of sample volumes, even routine application of recovery measurement, are not only discouraged, they would be directly counterproductive to a PWS's interest. That is, if one could only forget occurrences like Milwaukee, Carolton Ga, Rome NY, or Berlin NH.

From the standpoint of technical application, planning and implementation of sampling and analysis at any surface water location to define the true features of *Cryptosporidium* and *Giardia* concentrations is completely straightforward. Many PWSs water quality and watershed managers already know how it could be accomplished...outside the context of the LT2 regulation. However, if more efficient and effective monitoring were implemented the immediate result would be the appearance of 2–10 times higher organism concentrations

than previously apparent. As pointed out above the real exposure of the water consuming population would not have changed in any way, simply the monitoring perspective would be different. The new data would be unreconciled with previous estimates of risk based on previous data. Careful planning and a concerted efffort would also be required to ensure appropriate public perception of the new results.

What would be the benefits of following this path? First and foremost it would be a recognition of reality rather than propagating what in reality is a dangerous misconception, that is, the belief that zero's mean absence in the face of some real (even though perhaps objectively low) risk of waterborne outbreaks. Second, real (nonzero) data describing the concentrations of *Cryptosporidium* and *Giardia* over typical annual cycles will provide the individual PWS with knowledge of where the organisms are coming from, how they vary over time, when they are likely to be at highest levels enabling rational teatment system management, and ability to determine the effectiveness of management procedures. Finally, and not least of all, such data will reveal the true spectrum of occurrence from lowest to highest. From a review of waterborne outbreak occurrence it is not clear that the risk of an outbreak is directly proportional to the concentration of these organisms. Indeed most waterborne outbreak locations mentioned above have the benefit of relatively high quality sources and presumably relatively low *Cryptosporidium* and *Giardia* concentrations. Accurate concentration data will enable beginning to understand their relation to risk and their potentially useful relationship to other more easily monitored water quality parameters.

In terms of considering the wisdom of applying a second round of *Cryptosporidium* monitoring under LT2 questions remain for which answers could be found through additional preliminary investigation: (1) What is the impact of measuring recovery efficiency and calculation of concentration on the resulting position of a PWS within the concentration-risk spectrum? (2) Are recovery efficiency measurements reproducible and are their variations significant in describing *Cryptosporidium* and *Giardia* levels? (3) Can information from watershed surveys be used to indicate the approximate position within the risk spectrum? and 4) Do available data support four tiers of risk management covering the range of organism concentrations and variability?

The knowledge, understanding, and tools are available to resolve unanswered questions and to proceed toward rational implementation of procedures in the interest of efficient and effective public health protection. Ability to follow this course is at this stage will depend on far-sighted and creative interaction between the regulatory and water supply communities within a restrictive regulatory framework.

■ ASSOCIATED CONTENT

Supporting Information

Detail from USEPA public meeting presentations, Supporting Information Figures A and B, and full size copies of Figures 1, 5, 6a, b, and 8a–h, not practical to include in the body of the article, are provided here as Supporting Information. Supporting Information Figure A. USEPA tentative schedule of LT2ESWTR Regulatory Review Events.[7] Supporting Information Figure B. USEPA summary of LT2 *Cryptosporidium* occurrence in 39,700 field samples from 1670 water intake sampling sites used by 1376 U.S. PWS's serving >10 000 population.[8] This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: ++61 0408 548 427; fax: ++61 2 4221 3238; email: jongerth@uow.edu.au.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) USEPA. Long Term 2 Enhanced Surface Water Treatment Rule. 2006. https://federalregister.gov/a/06-4.

(2) USEPA. National primary drinking water regulation: Monitoring for public water supplies; Final Rule. *Fed. Regist.* **1996***61*(94):24353–24388

(3) USEPA. Occurrence and Exposure Assessment for the Final Long Term 2 Enhanced Surface Water Treatment Rule. Office of Water (4606-M), EPA 815-R-06-002 December 2005.

(4) Finn, M.. Overview of LT2 Rule Requirements. Presentation to 2nd Stakeholder Meeting, April 2012. http://water.epa.gov/lawsregs/rulesregs/sdwa/lt2/upload/LT2rulerequirements1.pdf.

(5) Ongerth, J. E. *Giardia* cyst concentrations in river water. *J. Am. Water Wks. Assoc.* **1989**, *81* (9), 81–86.

(6) Hansen, J.; Ongerth, J. E. Effects of time and watershed characteristics on the concentration of *Cryptosporidium* oocysts in river water. *Appl. Environ. Microbiol.* **1991**, *57* (10), 2790–2795.

(7) Miller, W. Regulatory Review Process for LT2, Presentation to 2nd LT2 Stakeholder Meeting, April 2012. http://water.epa.gov/lawsregs/rulesregs/sdwa/lt2/upload/regreviewprocess2.pdf.

(8) Messner, M. USEPA LT2 Round 1 *Cryptosporidium* Occurrence, 1st Stakeholder Meeting, December 2011. http://water.epa.gov/lawsregs/rulesregs/sdwa/lt2/upload/lt2round1crypto.pdf.

(9) Ongerth, J. E. *Cryptosporidium* Occurrence, Variations, and MS Recovery. Presentation to 3rd LT2 Stakeholder Meeting, November, 2012.

(10) Ongerth, J. E. The LT2 *Cryptosporidium* data...What do they tell us about *Cryptosporidium* in surface water in the USA? *Environ. Sci. Technol.* **2013a**, *47* (9), 4029–4038.

(11) USEPA,. Method 1623: Cryptosporidium and Giardia in Water by Filtration/IMS/FA. EPA 815-R-05-002. http://www.epa.gov/nerlcwww/documents/ 1623de05.pdf.

(12) Ongerth. Concentrations of *Cryptosporidium* and *Giardia* in water—The role and importance of recovery efficiency. *Water Res.* **2013b**, *47*, 2479–2488.

(13) Ongerth, J. E. *Evaluation of Flow Equalization in Wastewater Treatment*, EPA-600/2-79-096; Municipal Environmental Research Laboratory: Cincinnati, OH, 1979.

(14) Rose, J. B.; Gerba, C. P.; Jakubowski, W. Survey of potable water supplies for *Cryptosporidium* and *Giardia*. *Environ. Sci. Technol.* **1991**, *25* (8), 1393–1400.

(15) LeChevallier, M. W.; Norton, W. D.; Lee, R. G. Occurrence of *Giardia* and *Cryptosporidium* spp. in surface water supplies. *Appl. Environ. Microbiol.* **1991**, *57* (9), 2610–2616.

(16) Karanis, P.; Schonen, D.; Seitz, H. M. Distribution and removal of *Giardia* and *Cryptosporidium* in water supplies in Germany. *Water Sci. Technol.* **1997**, *37* (2), 9–18.

(17) Francy, D. S.; Simmons, O. D., III; Ware, M. W.; Granger, E. J.; Sobsey, M. D.; Schaefer, F. W., III Effects of seeding procedures and water quality on recovery of *Cryptosporidium* oocysts from stream water by using US Environmental Protection Agency Method 1623. *Appl. Environ. Microbiol.* **2004**, *70* (7), 4118–4128.

(18) Messner, M.; Wolpert, R. *Cryptosporidium* and *Giardia* occurrence in ICR drinking water sources—Statistical analysis of ICR Data. *Chapter 19 in Information Collection Rule Data Analysis*; McGuire, M., McLain, JL, Obolenski, A., Eds.; AWWA Research Foundation and American Water Works Association: Denver CO, 2002.

(19) Shaw P. K.; Brodsky R. E.; Lyman D. O.; Wood B. T.; Hibler C. P.; Healy G. R.; Macleod K. I.; Stahl W.; Schultz M. G. A communitywide outbreak of giardiasis with evidence of transmission by a municipal water supply. *Annals Of Internal Medicine [Ann Intern Med]*, IOct 1977; SSN: 0013-936X, Vol. *87* (4), pp 426−32

(20) Lippy, E. C. Tracing a giardiasis outbreak at Berlin, New Hampshire. *J. Am. Water Works Assoc.* **1978**, *70*, 512−520.

(21) U.S. Congress, 1996. Safe Drinking Water Act Amendments, Title I, Section 104 (c) (9). http://www.gpo.gov/fdsys/pkg/PLAW-104publ182/html/PLAW-104publ182.htm.