EPA Public Access

Author manuscript

Water Res. Author manuscript; available in PMC 2020 April 20.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Water Res. 2019 April 15; 153: 263–273. doi:10.1016/j.watres.2018.12.058.

Occurrence of Coliphage in Raw Wastewater and in Ambient Water: A Meta-analysis

Sharon P. Nappier^a, Tao Hong^b, Audrey Ichida^b, Ali Goldstone^b, Sorina E. Eftim^b

^aU.S. Environmental Protection Agency, Office of Water, Office of Science and Technology, 1200 Pennsylvania Avenue, NW, Washington, DC 20460, USA

^bICF, LLC, 9300 Lee Highway, Fairfax, VA 22031, USA

Abstract

Coliphage have been proposed as indicators of fecal contamination in recreational waters because they better reflect the persistence of pathogenic viruses in the environment and through wastewater treatment than traditional fecal indicator bacteria. Herein, we conducted a systematic literature search of peer-reviewed publications to identify coliphage density data (somatic and male-specific, or MSC) in raw wastewater and ambient waters. The literature review inclusion criteria included scope, study quality, and data availability. A non-parametric two-stage bootstrap analysis was used to estimate the coliphage distributions in raw wastewater and account for geographic region and season. Additionally, two statistical methodologies were explored for developing coliphage density distributions in ambient waters, to account for the nondetects in the datasets. In raw wastewater, the analysis resulted in seasonal density distributions of somatic coliphage (SC) (mean 6.5 log₁₀ plaque forming units (PFU)/L; 95% confidence interval (CI): 6.2–6.8) and MSC (mean 5.9 log₁₀ PFU/L; 95% CI: 5.5–6.1). In ambient waters, 49% of MSC samples were nondetects, compared with less than 5% for SC. Overall distributional estimates of ambient densities of coliphage were statistically higher for SC than for MSC (mean 3.4 and 1.0 log₁₀ PFU/L, respectively). Distributions of coliphage in raw wastewater and ambient water will be useful for future microbial risk assessments.

Keywords

raw wastewater; ambient water; systematic review; microbial risk assessment; male-specific coliphage; somatic coliphage

Introduction

Evidence from microbial risk assessments (Soller et al. 2010; Soller et al. 2015; Sunger et al. 2018) and epidemiology studies (Cabelli et al. 1982, Lee et al. 1997, Colford et al. 2005, Wiedenmann et al. 2006, Colford et al. 2007, Wade et al. 2010, Griffith et al. 2016) illustrates that viruses cause the majority of gastrointestinal illnesses associated with primary contact recreation in surface waters impacted by human sources. Further, U.S. outbreak

surveillance data point to viruses as the leading pathogen group responsible for ambient recreational water outbreaks (Jiang et al. 2007, Sinclair et al. 2009, Hlavsa 2015).

Human enteric viruses can enter recreational surface waters from both treated and untreated human sources, including treated wastewater effluent. Often wastewater treatment and associated disinfection processes specifically target the removal and inactivation of bacteria, but not viruses (U.S. EPA 1986, 2015). Numerous studies have identified the presence of viruses in wastewater treatment effluent, often when traditional fecal indicator bacteria (FIB) are non-detectable (da Silva et al. 2007, Haramoto et al. 2007, Kageyama et al. 2003, Kitajima et al. 2009, Kuo et al. 2010, Simmons and Xagorarakis 2011).

Coliphage are a subset of bacteriophage viruses that infect *Escherichia coli*. In particular, male-specific (MSC, also known as “F-specific”) and somatic coliphage (SC) have been proposed as more reliable indicators of human viral fecal contamination than traditional FIB in recreational waters (U.S. EPA 2015, Cabelli et al. 1982, Gerba 1987, Havelaar et al. 1993, Palmateer et al. 1991). Both coliphage exhibit greater similarity to human enteric viruses in their physical structure, composition, morphology, survivability in the environment, and persistence in treatment processes compared to traditional FIB (U.S. EPA 2015, Havelaar et al. 1993, Funderburg and Sorber 1985, Gantzer et al. 1998, Grabow 2004, Nappier et al. 2006, Pouillot et al. 2015). Coliphage are also useful in that they can be detected and quantified by simple, inexpensive, rapid, and reliable methods (U.S. EPA 2015, Gerba 1987, Havelaar 1987, Love and Sobsey 2007, Muniesa et al. 2018), and they are abundant in municipal wastewater and polluted waters (Debartolomeis and Cabelli 1991, Havelaar et al. 1990, Leclerc et al. 2000, Mandilara et al. 2006). They originate almost exclusively from the feces of humans and other warm-blooded animals, and undergo only limited replication in sewage under specific conditions, such as high densities of coliphage and susceptible host *E. coli* at permissive temperatures (Grabow 2004, Sobsey et al. 1995).

Previously, a systematic review methodology was developed to evaluate densities of key viruses in raw wastewater and published distributions of norovirus densities in raw wastewater (Eftim et al. 2017). These distributions are specifically useful for conducting quantitative microbial risk assessments (QMRA) (Soller et al. 2018; Nappier et al. 2018). For the derivation of coliphage-based recreational water quality criteria values, similar distributions of MSC and SC in raw wastewater and ambient waters are needed. However, to date, coliphage densities in the literature have only been summarized in influent for MSC (Pouillot et al. 2015) or were not modeled as distributions (McMinn et al. 2017).

The objectives of this work are: 1) to summarize the results of a systematic literature review; 2) to characterize the density distributions of MSC and SC in raw wastewater and in ambient waters; and 3) to identify if coliphage densities vary significantly by geographic region and season. The methodologies developed for this work will also offer utility to other QMRA applications, such as risk characterizations for consumption of recycled wastewater and shellfish, and exposure to recreational waters and biosolids.

2. Materials and Methods

2.1 Data sources and literature search

A systematic literature search of the peer reviewed literature for articles reporting coliphage densities in raw wastewater and ambient water in PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) and Web of Science was performed. The search included the keywords (somatic coliphage OR F-specific OR male-specific coliphage OR male-specific bacteriophage OR FRNA OR SOMCPH OR male coliphage) AND (water OR influent OR effluent OR ambient OR sewage OR wastewater). The literature search was limited to peer-reviewed publications written in English between 1990 and January 2017.

Study inclusion criteria were applied to each publication, which comprised scope (Step 1), study quality (Step 2), and data availability (Step 3). For Step 1, two independent screeners first evaluated the titles and abstracts of retrieved articles to assess whether the paper would likely include coliphage density data in ambient water, raw wastewater, or wastewater treatment effluent. Two reviewers then reviewed the full text for papers whose abstracts passed the first screening step to confirm that the paper contained quantitative density data for coliphage in ambient water and raw wastewater. Papers with only presence/absence data and papers reporting on animal manure wastewater facilities were excluded.

The Step 2 study quality screen included an evaluation of the assay type and clarity. Coliphage enumeration methods considered acceptable included: International Organization for Standardization (ISO) 10705–1 (ISO 1995); ISO 10705–2 (ISO 2000); United States Environmental Protection Agency (EPA) Method 1602 (U.S. EPA 2001a); EPA Method 1601 (U.S. EPA 2001b); and Standard Methods 9224 B, C or D (Eaton et al. 2005). When one of the aforementioned methods was cited, it was assumed the recommended host (e.g., CN-13 or Famp) was employed, unless an alternative host was specifically cited. It was also assumed controls, detection limits, and inhibition evaluation were all handled properly.

In addition to the methods mentioned above, papers containing coliphage data obtained using the methods cited in: Adams (1959), Debartolomeis and Cabelli (1991), EasyPhage method (Luther and Fujioka 2004), Havelaar and Hogeboom (1984) and U.S. EPA (1984) were also included. However, for these citations, use of one of the following host cell lines was a requirement: *E. coli* CN13 (also CN13) (ATCC 700609); *E. coli* HS(pFamp)R (ATCC 700891); *E. coli* WG5 (also WG5) (ATCC 700078) (also called *E. coli* CN (per ISO-10705–2)); *S. enteritidis* var Typhimurium WG49 (also WG49); *E. coli* ATCC 13706 (also strain C) (previously WG4); and *E. coli* K12 Hfr. If an alternative host cell line was cited (i.e., not one of the listed eligible hosts), the study was excluded from the review.

Coliphage density data enumerated by molecular methods was not included. Publications were also not included if the study and associated methodologies were not documented clearly. Additional information was collected for studies that passed the quality step including: specific water type (for ambient, marine, and freshwater), geographic region, sample size, and seasonal information (or sampling dates).

After a publication passed both the scope and quality criteria, it was evaluated for data availability (Step 3). For this screening step, a requirement was for publications to include individual data points, not summary measures (i.e., means or medians). In some cases, individual data points were available in a table or were digitized from figures using GetData Graph Digitizer, Version 2.25 (<http://www.getdata-graph-digitizer.com/>) and WebPlotDigitizer Version 3.12 (<http://arohatgi.info/WebPlotDigitizer/app/>). Data from study authors was requested when raw data were not available in the manuscript, figures, or supplemental materials. If the author sent data, and those data aligned with the published information, those datasets were included. For example, Pouillot et al. (2015) used data collected by the Food and Drug Administration (FDA), and FDA provided the individual data points collected for that study. If a research group published multiple papers using the same dataset, only one publication was used to represent the dataset.

2.2 Statistical approaches

Coliphage samples were collected under a variety of conditions and enumerated using different analytical methods in the studies selected. Several studies quantified both MSC and SC from the same sample. In such cases, a sample provided two individual data points for our analysis (one for MSC and one for somatic). For this meta-analysis, each individual data point was characterized by the month of sample collection, the coliphage type (i.e., MSC or somatic), and the quantified viral density in either raw wastewater or ambient water. Viral density units are reported in log₁₀ plaque forming units per liter (PFU/L). Data extraction and management was conducted using Excel (Microsoft Corporation 2013). Statistical analyses were performed with R version 3.3.3 (R Development Core Team 2013).

Descriptive statistics (mean, median, and standard deviation) were calculated for the observed MSC and SC density data in both raw wastewater and ambient waters. Normal Q-Q plots, Shapiro-Wilk test, and Kolmogorov-Smirnov (K-S) test were used to evaluate normality of the data. When the data were normal, the observed mean log densities of coliphage were compared by geographic region and by season, using two sample t-tests (MSC compared to SC, for each season/region), with the assumption of unequal variances. Log means of coliphage were compared across region (or season) using ANOVA. When the ANOVA test rejects the null hypothesis of no differences, we used Tukey's post-hoc honest significance test (Yandell 1997), which allows for multiple comparisons to evaluate regional and seasonal differences. When data were not normal (or when the assumption of equal variances under ANOVA was not satisfied), we performed the corresponding non-parametric Mann-Whitney and Kruskal-Wallis tests, and we used Dunn's multiple nonparametric pairwise tests (Dunn 1964) following rejection of a Kruskal-Wallis test.

Hypothesis tests were not conducted for simultaneous comparisons. For all tests, we assumed a significance level of 0.05.

The geographic regions included North America, Europe, Asia, Africa, South America, and Oceania. For countries in the Northern Hemisphere, the following seasonal definitions are used: spring (March to May), summer (June to August), fall (September to November), and winter (December to February) (source: <http://www.timeanddate.com/calendar/aboutseasons.html>). For South Africa, South America, and Oceania, the following seasonal

definitions are used: spring (September to November), summer (December to February), fall (March to May), and winter (June to August).

2.2.1 Coliphage distributions in raw wastewater—To develop coliphage distributions in raw wastewater, a previously published approach was applied to evaluate densities of key viral pathogens in raw sewage (Eftim et al. 2017). Briefly, a two-stage bootstrap analysis was applied to datasets for each coliphage type (Efron and Tibshirani 1994, Hastie et al. 2001). In Stage 1, the dataset was first separated into subsets based on geographic region and season. Samples sizes equal to the original region- or season-specific subsets were drawn with replacement and with equal weight, for 10,000 iterations, resulting in a distribution of estimated means, standard deviations and 95% confidence intervals. The objective of this stage was to understand variation in the distribution of coliphage densities associated with geographic region and season (Table 2).

In Stage 2, a stratified bootstrap approach incorporated variation contributed by each subset (geographic region and season). Unlike in Stage 1, the number of samples randomly selected from each subset were allowed to vary and were proportional to the subset sample size. Thus, the approach accounts for the large variability in the number of coliphage samples taken under different seasons and across geographic regions. For the Stage 2 analysis, the sampling step was repeated for 100, 500, 1,000, and 3,000 iterations. The samples were pooled and the means, standard deviations and their 95% confidence intervals were estimated. The means of the bootstrapped distributions were compared to each other using the Kolmogorov-Smirnov test (Lehmann and D'Abrera 2006).

2.2.2 Coliphage distributions in ambient waters—To develop distributions of coliphage densities in ambient waters, an approach that accounts for a potentially large number of values below the detection limit, often called left censored data, was employed. Several approaches can be used to develop descriptive statistics with censored data in the environmental field including: substitution, maximum likelihood, nonparametric methods, and semi-parametric methods. Typically, the amount of available uncensored data determines the best-suited approach (Helsel 2005).

First, using boxplots and summary statistics, the variability in the detected data only was evaluated. Next, two statistical approaches to estimate the distribution that best fit the data, accounting for the nondetects: 1) a non-parametric (Kaplan-Meier) method, which requires no assumption about a particular distributional shape, and 2) the maximum likelihood approach, which assumes the distribution (normal, lognormal, or other form) will closely fit the shape of the observed data was used. When censoring does not occur, the two approaches provide identical results. We compared the best fit under three separate approaches: when using the detected data only, the Kaplan-Meier approach, and the maximum likelihood approach. The ambient data distributions were also graphically compared using the cumulative distribution functions of the empirical distribution function and the best fit theoretical distribution. We compared the distributions using the two sample Kolmogorov-Smirnov test.

3. Results

Figure 1 illustrates the overall literature search strategy and number of titles retrieved. A total of 459 titles remained after duplicates from the two databases were removed. Literature Search 1 was conducted in July 2014, and it was updated with Literature Searches 2 and 3, which were conducted in September 2015 and January 2017, respectively. After a full-text review, 104 references passed study scope inclusion criteria for raw wastewater and 122 references passed study scope inclusion criteria for ambient water. Of the studies that passed scope, 79 passed study quality inclusion criteria for raw wastewater, and 94 passed study quality inclusion criteria for ambient water. Individual data points were obtained for 17 studies for raw wastewater (Table 1) and for 21 studies for ambient water (Table 2).

Together these studies provided a total of 2,124 individual data points (Table 1, coliphage in raw wastewater, $n=333$ [somatic] and $n=397$ [MSC]; Table 2, coliphage in ambient water, $n=391$ [somatic] and $n=1,011$ [MSC]). In raw wastewater, only one data point was reported below LOD, which was substituted with the LOD reported in the publication (Table 1, Blanch et al. 2006). However, in ambient waters, 5% of the SC and 49% of the MSC data were below the LOD (Table 2). The literature searches yielded references with quantitative data representing six geographic regions, including 20 countries (for both raw wastewater and ambient (see Tables 1 and 2).

3.1 Descriptive Statistics

Table 3 presents the summary statistics of the observed coliphage densities in raw wastewaters. For MSC and SC in raw wastewater, the observed densities ranged from 1.7 to 8.34, and from 2.0 to 7.98 $\log_{10}$ PFU/L, respectively (data not shown). Approximately 82% of all samples were collected in North America and Europe, with the remaining 18% from China, Japan, Singapore, and South Africa (Table 3). The Shapiro-Wilk and Kolmogorov-Smirnov tests for normality rejected the null hypotheses of normality for the season and geographic region-specific observed MSC and SC densities in raw wastewater (K-S test, p -values < 0.001), except for MSC densities in Asia, SC densities in North America and Africa, and SC in the fall. Observed SC densities were higher in fall, winter and spring compared to summer (Kruskal-Wallis test, p -value = 0.006), and observed MSC densities were higher in winter and fall compared to spring and summer (Kruskal-Wallis test, p -value < 0.001) (Figure 2B). However, it should be noted that seasonal information was unavailable for approximately 32% of the data.

The observed median densities of SC in raw wastewater were consistently significantly higher compared to median densities of MSC, a phenomenon independent of geographic region (Figure 2A) and season (Figure 2B). The exception was in North America, where the median density of MSC was significantly higher than median for SC (Mann-Whitney test, p -value = 0.002).

Table 4 presents the summary statistics of the observed coliphage densities in ambient waters. The observed densities for the MSC and SC in ambient waters ranged from -0.52 to 6.85, and from 0.95 to 6.96 $\log_{10}$ PFU/L, respectively (data not shown). Approximately 60% of all ambient samples were collected in North America and Europe, with the remaining

40% from South America, Asia and Oceania (Table 4). The Shapiro-Wilk tests for normality revealed that the season and geographic region-specific observed MSC and SC ambient densities were normally distributed (Shapiro-Wilk test, p -values < 0.001), except for SC in Asia, both MSC and SC in Oceania, MSC in spring and fall, and SC in fall and winter. The SC ambient densities were significantly different across the geographic regions (Kruskal-Wallis test, p -values < 0.05) (Figure 2C). MSC mean ambient densities were significantly different between North America and Europe, and Asia, and between Europe and South America and Asia (multiple comparison test after Kruskal-Wallis tests, p -value < 0.05). The lowest observed mean MSC ambient densities were in Oceania, and the highest were observed in Asia (Table 4).

In terms of seasonality, observed mean SC ambient densities were higher in spring and summer compared to fall and winter (Figure 2D) (summer *vs* winter t-test, p -value < 0.001). Additionally, the observed median ambient densities of SC were consistently significantly higher compared to median ambient densities of MSC in Europe, Oceania, spring and summer (p -values for MSC *versus* SC comparisons Mann Whitney test < 0.001). Although the same trend is present in North America and in the fall, the difference is not significant (Mann Whitney test, p -values > 0.05). For Asia and in the winter the median ambient densities of SC are lower than for MSC, but not significantly so.

3.2 Distributions of coliphage densities in raw wastewater

Next, a two-stage bootstrap analysis was conducted to further understand variation in the distribution of coliphage densities associated with geographic region and season. Table 5 presents the results of the Stage 1 bootstrap analysis. For MSC, there is a significant difference in the mean density distribution between Europe and North America and Asia (Figure 3, left panel) (Dunn's test, p -values < 0.01), and for somatic, there is a significant difference in the mean density distributions between all geographic locations (Dunn's test, p -value < 0.01) (Figure 3, right panel). Consistently across geographic regions and across seasons, the mean SC densities are higher than the mean MSC densities (t-test, p -values < 0.001). The only exception is in North America where the SC mean density is statistically significantly lower than the mean density of MSC ($5.2 \log_{10}$ PFU/L versus $5.9 \log_{10}$ PFU/L, t-test, p -value < 0.001) (Table 5).

A seasonal difference in mean coliphage density in raw wastewater was also observed, with the greatest differences between spring and summer for both MSC and SC (Dunn's test, p -values < 0.01), with a winter peak for both coliphage types (Table 5 and Figure 4). The season-specific bootstrapped distributions obtained in Stage 1 are all significantly different from each other for both MSC and SC (Kruskal-Wallis test p -value < 0.001 for comparisons among seasons).

However, application of the Stage 2 bootstrap analysis (Table 5), which accounts for the number of samples taken from each geographic region and season sample subset, resulted in no statistically significant difference between the seasonal and geographic region distributional estimates created for MSC (means 5.8 and $5.9 \log_{10}$ PFU/L, respectively; t-test, p -values < 0.001) and SC (means 6.0 and $6.5 \log_{10}$ PFU/L, respectively; t-test, p -values < 0.001). For the Stage 2 bootstrap analysis, a sensitivity analysis additionally

evaluated the impact of the number of sampling iterations on the bootstrapped coliphage density distributions (for 100, 500, 1,000, and 3,000 iterations), and indicated the number of iterations did not affect results (Figures S1 and S2 in Supplemental File). The results for 3,000 iterations in this analysis are presented (Table 5).

3.3 Distributions of coliphage densities in ambient waters

Table 6 compares the estimated means and standard deviations of the distributions of coliphage in ambient waters using: 1) the detected data only (i.e., excluding the nondetects); 2) the Kaplan-Meier non-parametric method; and 3) the maximum likelihood approach, with the assumption of normality on the entire dataset. The results illustrate the three methods are comparable when the number of nondetects in the dataset is small (<5%), which is the case for SC. However, for the MSC dataset, 49% of the samples are nondetects. The mean of the distribution of MSC in ambient waters estimated when excluding the nondetects is significantly higher than the means estimated including nondetects (2.2 versus 1.6 and 1.4 log₁₀ PFU/L, t-test p-values <0.001) (Table 6). The Kaplan-Meier and the maximum likelihood approaches yield similar results for MSC and for SC. The best-fit parameters for the distributions of both MSC and SC in ambient waters are provided in Table 6. For SC, the Kaplan-Meier and MLE distribution are not statistically different (K-S test, p-value >0.05), while the empirical distribution based on the detected data, under predicts slightly (K-S test p-value =0.014). For MSC, the empirical distribution CDF (cumulative distribution function) lies under the Kaplan-Meier and the MLE distributions. Additionally, the theoretical and empirical probability distribution curves of the data are visually discordant due to the large percent of nondetects of MSC in ambient waters (Figure S3).

4. Discussion

This work summarizes coliphage density data in raw wastewater and in ambient waters. Distributions of indicators and pathogens based on a critical review of the literature add value to risk assessments (Soller et al. 2017, 2018) and a greater understanding water and wastewater treatment needs (McMinn et al. 2017, Pouillot et al. 2015, Olivieri et al. 2016). We applied a systematic literature search approach in identifying relevant literature. The study selection process is reproducible and can be modified to incorporate additional data, as they become available.

Developing distributions of pathogen densities in various environmental matrices can be challenging, with respect to data availability and the amount of data below the detection limits (Davies et al. 2003, Till et al. 2008, DiDonato et al. 2009). In addition, data are heterogeneous with respect to geographical region and sampling season, and not all authors report a sampling date or season. In this work, computational approaches tailored to the type and amount of data available are presented. Distributions were developed using a meta-analysis approach, combining heterogeneous data from all the available studies, which provides more comprehensive information than individual study-based distributions (Eftim et al. 2017).

When few data points were below the LOD, the two-step bootstrap analysis produced a distribution mean and standard deviation. The bootstrap analysis is useful because, it: 1) is

non-parametric; 2) does not assume a shape for the underlying statistical distribution; and 3) allows for the evaluation of and the accounting for subsets of data of various sizes, sampled at different geographic regions and over different seasons. Seasonal and geographic region influences might be especially important for certain pathogens (Grassley and Fraser 2006, Gerba et al. 2017, Cherrie et al. 2018), such as norovirus, which is found in higher densities in the winter and spring (Eftim et al. 2017).

For left-censored data, the Kaplan-Meier or maximum likelihood approaches can be useful for limiting uncertainties in distribution estimates (Helsel, 2005). Lim et al. (2015) used a similarly left-censored data maximum likelihood regression technique to estimate parameters that characterize the viral distributions for norovirus and adenovirus in surface waters. Our analysis shows that accounting for nondetects is important, especially when approximately half of the dataset includes non-detects (49%). Instead of applying a commonly used strategy of substituting nondetects with preset values (i.e., half the limit of detection), which has been shown to yield biased distributional parameter estimates (Helsel 2005), we employed multiple statistical methods for left-censored data. Our results indicate that excluding the nondetects yields distributions with very different means compared with when including nondetects, especially for MSC (Table 6).

Further, because the Kaplan-Meier and maximum likelihood results are very similar for MSC, a goodness of fit of the log-normal shape of the distribution to the data is suggested. The log-normal assumption is consistent with the finding that most microbial data in the environment are log-normally distributed (Esmen and Hammad 1977, Wymer et al. 2007). However, in situations where the shape of the distribution is not known or cannot be determined based on the detected data only (e.g., when the sample size is small), the nonparametric Kaplan-Meier method may be an appropriate way to estimate distributional parameters (Helsel 2005).

Our study results illustrate that observed SC densities significantly outnumber MSC densities in both raw wastewater and in ambient waters (Tables 3 and 4), which is consistent with the literature (Moce-Llivina et al. 2005, DiDonato et al. 2009, McMinn et al. 2017). Additionally, in ambient waters, SC are detected in a higher percentage of samples, as compared to MSC (Table 6). The estimated distributions of our two-stage bootstrap analysis also result in distributions with a higher density of SC than MSC in raw wastewater (overall means 6.0 and 5.8 $\log_{10}$ PFU/L, respectively, Table 5), but the difference between the two coliphage types is not significant when the analysis accounts for geographic region or season (Table 5).

These bootstrapped estimates for MSC and SC in raw wastewater are similar to meta-analysis estimates previously reported in the literature. McMinn et al. (2017) reported observed mean densities of MSC and SC in raw wastewaters (5.2 and 5.3 $\log_{10}$ PFU/L, respectively). Additionally, in an earlier meta-analysis, Pouillot et al. (2015) reported a conditional posterior distribution of MSC in raw wastewater with a mean of 6.2 $\log_{10}$ PFU/L. While results are overall similar, several reasons for the differences in our study include: different study inclusion criteria, different numerical evaluation approaches, and

due to the timing of our analysis, inclusion of data from several newly published studies (e.g., Bailey et al. 2017, Montazeri et al. 2015).

A seasonal difference in observed mean coliphage densities was observed in raw wastewater and in ambient waters. For raw wastewater, there was a winter peak for both coliphage types in the bootstrap analysis. This seasonal variability in MSC and SC is similar to what was observed for norovirus in raw wastewater, which had significantly higher peaks in winter and spring compared to summer and fall (Eftim et al. 2017). The winter peak in MSC is also consistent with Pouillot et al. (2015), where authors noted a significant increase in predicted mean densities of MSC between February to June (i.e., winter and spring), as compared to their twelve-month mean density estimate.

5. Conclusions

The systematic literature search and a two-stage bootstrap analysis provide robust methods for developing distributions for indicator and pathogen densities. Additionally, the Kaplan-Meier or maximum likelihood approaches can be useful for limiting uncertainties in distributional estimates, when datasets contain nondetects. The resulting distributions of MSC and SC in raw wastewater and ambient waters will be useful in the future for a wide range of risk assessment applications in diverse geographic regions, including water reuse, shellfish, and understanding recreational risks from epidemiological data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The research described in this article was funded by the U.S. EPA Office of Water, Office of Science and Technology. This work has been subject to formal Agency review. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. EPA. Thank you to Pouillot and co-authors at the FDA for providing EPA with the dataset used in their analysis. We are grateful to Dr. Alexandria Boehm for providing expertise on the standard methods and host strains, and for review our draft manuscript. We thank Isaac Warren and Jennifer Hsieh for help in literature review, figure digitization, and data extraction. We are also grateful for authors who shared their full datasets including Drs. Tiong Gim Aw, Anicet Blanch, Mark Sobsey, and Pierre Payment. We are grateful to Brian Schnitker, Dr. Shamima Akhter, and Dr. Jamie Strong for their invaluable review and technical edit of our draft manuscript.

References

- Adams MH (1959) Bacteriophages, Interscience Publishers, New York, NY.
- Astrom J, Pettersson TJ, Stenstrom TA and Bergstedt O (2009) Variability analysis of pathogen and indicator loads from urban sewer systems along a river. *Water Science and Technology* 59(2), 203–212. [PubMed: 19182329]
- Aw TG and Gin KY (2010) Environmental surveillance and molecular characterization of human enteric viruses in tropical urban wastewaters. *Journal of Applied Microbiology* 109(2), 716–730. [PubMed: 20233263]
- Bailey ES, Price M, Casanova LM and Sobsey MD (2017) E. coli CB390: An alternative E. coli host for simultaneous detection of somatic and F+ coliphage viruses in reclaimed and other waters. *Journal of Virological Methods* 250, 25–28. [PubMed: 28939117]
- Blanch AR, Belanche-Munoz L, Bonjoch X, Ebdon J, Gantzer C, Lucena F, Ottoson J, Kourtis C, Iversen A, Kuhn I, Moce L, Muniesa M, Schwartzbrod J, Skrabber S, Papageorgiou GT, Taylor H,

- Wallis J and Jofre J (2006) Integrated analysis of established and novel microbial and chemical methods for microbial source tracking. *Applied and Environmental Microbiology* 72(9), 5915–5926. [PubMed: 16957211]
- Brezina SS and Baldini MD (2008) Detection of somatic coliphages as indicators of faecal contamination in estuarine waters. *Revista Argentina de Microbiología* 40(1), 72–74. [PubMed: 18669057]
- Cabelli VJ, Dufour AP, McCabe LJ and Levin MA (1982) Swimming-associated gastroenteritis and water quality. *American Journal of Epidemiology* 115, 606–616. [PubMed: 7072706]
- Carducci A and Verani M (2013) Effects of bacterial, chemical, physical and meteorological variables on virus removal by a wastewater treatment plant. *Food and Environmental Virology* 5(1), 69–76. [PubMed: 23412727]
- Carducci A, Verani M, Battistini R, Pizzi F, Rovini E, Andreoli E and Casini B (2006) Epidemiological surveillance of human enteric viruses by monitoring of different environmental matrices. *Water Science and Technology* 54(3), 239–244.
- Cherrie MPC, Nichols G, Iacono GL, Sarran C, Hajat S and Fleming LE (2018) Pathogen seasonality and links with weather in England and Wales: a big data time series analysis. *BMC Public Health* 18(1), 1067. [PubMed: 30153803]
- Colford JM Jr., Wade TJ, Schiff KC, Wright CC, Griffith JF, Sandhu SK, Burns S, Sobsey M, Lovelace G and Weisberg SB (2007) Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. *Epidemiology* 18(1), 27–35. [PubMed: 17149140]
- Colford JM, Wade TJ, Schiff KC, Wright C, Griffith JF, Sandhu SK and Weisberg SB (2005) Recreational water contact and illness in Mission Bay, California, Southern California Coastal Water Research Project.
- da Silva AK, Le Saux JC, Parnaudeau S, Pommepuy M, Elimelech M and Le Guyader FS (2007) Evaluation of removal of noroviruses during wastewater treatment, using real-time reverse transcription-PCR: different behaviors of genogroups I and II. *Applied and Environmental Microbiology* 73(24), 7891–7897. [PubMed: 17933913]
- Davies CM, Yousefi Z and Bavor HJ (2003) Occurrence of coliphages in urban stormwater and their fate in stormwater management systems. *Letters in Applied Microbiology* 37(4), 299–303. [PubMed: 12969492]
- Debartolomeis J and Cabelli VJ (1991) Evaluation of an *Escherichia coli* host strain for enumeration of F male-specific bacteriophages. *Applied and Environmental Microbiology* 57(5), 1301–1305. [PubMed: 1830197]
- Delignette-Muller ML and Dutang C (2015) Fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software* 64(4).
- DiDonato GT, Stewart JR, Sanger DM, Robinson BJ, Thompson BC, Holland AF and Van Dolah RF (2009) Effects of changing land use on the microbial water quality of tidal creeks. *Marine Pollution Bulletin* 58(1), 97–106. [PubMed: 18922549]
- Dunn OJ (1964) Multiple comparisons using rank sums. *Technometrics* 6, 241–252.
- Eaton AD, Franson MAH, eds. (2005) *Standard Methods for the Examination of Water and Wastewater*. Washington, DC: APHA, AWWA, WEF.
- Efron B and Tibshirani RJ (1994) *An Introduction to the Bootstrap*. London: Chapman & Hall/ CRC Press.
- Eftim SE, Hong T, Soller J, Boehm A, Warren I, Ichida A and Nappier SP (2017) Occurrence of norovirus in raw sewage - A systematic literature review and meta-analysis. *Water Research* 111, 366–374. [PubMed: 28110140]
- Esmen NA and Hammad YY (1977) Log-normality of environmental sampling data. *Journal of Environmental Sciences, Part A* 12(1–2), 29–41.
- Francy DS, Stelzer EA, Bushon RN, Brady AM, Williston AG, Riddell KR, Borchardt MA, Spencer SK and Gellner TM (2012) Comparative effectiveness of membrane bioreactors, conventional secondary treatment, and chlorine and UV disinfection to remove microorganisms from municipal wastewaters. *Water Research* 46(13), 4164–4178. [PubMed: 22682268]
- Funderburg SW and Sorber CA (1985) Coliphages as indicators of enteric viruses in activated sludge. *Water Research* 19(5), 547–555.

- Gantzer C, Maul A, Audic JM and Schwartzbrod L (1998) Detection of infectious enteroviruses, enterovirus genomes, somatic coliphages, and *Bacteroides fragilis* phages in treated wastewater. *Applied and Environmental Microbiology* 64(11), 4307–4312. [PubMed: 9797281]
- Gerba CP (1987). Phage as indicator of fecal pollution In: *Phage ecology*, (Goyal SM, Gerba CP, Bitton G, eds). New York, NY: John Wiley and Sons.
- Gerba CP, Betancourt WQ and Kitajima M (2017) How much reduction of virus is needed for recycled water: a continuous changing need for assessment? *Water Research* 108, 25–31. [PubMed: 27838026]
- Gourmelon M, Caprais MP, Mieszkis S, Marti R, Wery N, Jarde E, Derrien M, Jadas-Hecart A, Communal PY, Jaffrezic A and Pourcher AM (2010) Development of microbial and chemical MST tools to identify the origin of the faecal pollution in bathing and shellfish harvesting waters in France. *Water Research* 44(16), 4812–4824. [PubMed: 20709349]
- Gourmelon M, Caprais MP, Segura R, Le Mennec C, Lozach S, Piriou JY and Rince A (2007) Evaluation of two library-independent microbial source tracking methods to identify sources of fecal contamination in French estuaries. *Applied and Environmental Microbiology* 73(15), 4857–4866. [PubMed: 17557850]
- Grabow W (2004) Bacteriophages: update on application as models for viruses in water.
- Grassley NC and Fraser C (2006) Seasonal infectious disease epidemiology. *Proceedings of the Royal Society B: Biological Sciences*. 273(1600), 2541–2550. [PubMed: 16959647]
- Griffith JF, Weisberg SB, Arnold BF, Cao Y, Schiff KC and Colford JM Jr. (2016) Epidemiologic evaluation of multiple alternate microbial water quality monitoring indicators at three California beaches. *Water Research* 94, 371–381. [PubMed: 27040577]
- Haramoto E, Katayama H, Oguma K and Ohgaki S (2007) Quantitative analysis of human enteric adenoviruses in aquatic environments. *Journal of Applied Microbiology* 103(6), 2153–2159. [PubMed: 18045398]
- Haramoto E, Katayama H, Oguma K, Yamashita H, Tajima A, Nakajima H and Ohgaki S (2006) Seasonal profiles of human noroviruses and indicator bacteria in a wastewater treatment plant in Tokyo, Japan. *Water Science and Technology* 54(11–12), 301–308. [PubMed: 17302333]
- Haramoto E, Otagiri M, Morita H and Kitajima M (2012) Genogroup distribution of F-specific coliphages in wastewater and river water in the Kofu basin in Japan. *Letters in Applied Microbiology* 54(4), 367–373. [PubMed: 22324376]
- Haramoto E, Kitajima M, Katayama H, Asami M, Akiba M, Kunikane S (2009). Application of real-time PCR assays to genotyping of F-specific phages in river water and sediments in Japan. *Water Research* 43, 3759–3764. [PubMed: 19555992]
- Hastie T, Tibshirani R and Friedman JH (2001) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.
- Havelaar AH and Hogeboom WM (1984) A method for the enumeration of male-specific bacteriophages in sewage. *Journal of Applied Bacteriology* 56(3), 439–447. [PubMed: 6378873]
- Havelaar AH (1987) Bacteriophages as model organisms in water treatment. *Microbiology Science* 4(12), 362–364.
- Havelaar AH, Pot-Hogeboom WM, Furuse K, Pot R and Hormann MP (1990) F-specific RNA bacteriophages and sensitive host strains in faeces and wastewater of human and animal origin. *Journal of Applied Bacteriology* 69(1), 30–37. [PubMed: 2204615]
- Havelaar AH, van Olphen M and Drost YC (1993) F-specific RNA bacteriophages are adequate model organisms for enteric viruses in fresh water. *Applied and Environmental Microbiology* 59(9), 2956–2962. [PubMed: 8215367]
- Helsel DR (2005) *Nondetects and data analysis: statistics for censored environmental data*, Wiley-Interscience.
- Hlavsa M, Roberts V, Kahler A, Hilborn E, Mecher T, Beach M, Wade TJ, Yoder JS (2015) Outbreaks of Illness Associated with Recreational Water—United States, 2011–2012. *Mortality and Morbidity Weekly Report*. Center for Disease Control, 64(24):668–72.
- ISO (1995) *Water quality -- Detection and enumeration of bacteriophages -- Part 1: Enumeration of F-specific RNA bacteriophages ISO 10705-1*, International Organization for Standardization, Geneva, Switzerland.

- ISO (2000) Water quality -- Detection and enumeration of bacteriophages -- Part 2: Enumeration of somatic coliphages. ISO 10705-2, Geneva, Switzerland.
- Jebri S, Muniesa M and Jofre J (2017) General and host-associated bacteriophage indicators of fecal pollution In: Rose JB and Jiménez-Cisneros B, (eds) Global Water Pathogens Project. <http://www.waterpathogens.org> (Farnleitner A, and Blanch A (eds) Part 2 Indicators and Microbial Source Tracking Markers) <http://www.waterpathogens.org/book/coliphage> Michigan State University, E. Lansing, MI, UNESCO.
- Jiang SC and Chu W (2004) PCR detection of pathogenic viruses in southern California urban rivers. *Journal of Applied Microbiology* 97(1), 17–28. [PubMed: 15186438]
- Jiang SC, Chu W and He JW (2007) Seasonal detection of human viruses and coliphage in Newport Bay, California. *Applied and Environmental Microbiology* 73(20), 6468–6474. [PubMed: 17720839]
- Kageyama T, Kojima S, Shinohara M, Uchida K, Fukushi S, Hoshino FB, Takeda N and Katayama K (2003) Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *Journal of Clinical Microbiology* 41(4), 1548–1557. [PubMed: 12682144]
- Kitajima M, Haramoto E, Phanuwat C, Katayama H and Ohgaki S (2009) Detection of genogroup IV norovirus in wastewater and river water in Japan. *Letters in Applied Microbiology* 49(5), 655–658. [PubMed: 19780954]
- Kuo DH, Simmons FJ, Blair S, Hart E, Rose JB and Xagorarakis I (2010) Assessment of human adenovirus removal in a full-scale membrane bioreactor treating municipal wastewater. *Water Research* 44(5), 1520–1530. [PubMed: 19944439]
- Leclerc H, Edberg S, Pierzo V and Delattre JM (2000) Bacteriophages as indicators of enteric viruses and public health risk in groundwaters. *Journal of Applied Microbiology* 88(1), 5–21. [PubMed: 10735238]
- Lee JE, Lee H, Cho YH, Hur HG and Ko G (2011) F+ RNA coliphage-based microbial source tracking in Water Researchources of South Korea. *Science of The Total Environment* 412–413, 127–131.
- Lee JV, Dawson SR, Ward S, Surman SB and Neal KR (1997) Bacteriophages are a better indicator of illness rates than bacteria amongst users of a white water course fed by a lowland river. *Water Science and Technology* 35, 165–170.
- Lehmann EL and D'Abrera HJM (2006) *Nonparametrics: statistical methods based on ranks*. Springer, New York.
- Lim KY, Hamilton AJ and Jiang SC (2015) Assessment of public health risk associated with viral contamination in harvested urban stormwater for domestic applications. *Science of The Total Environment* 523, 95–108. [PubMed: 25863500]
- Locas A, Martinez V and Payment P (2010) Removal of human enteric viruses and indicator microorganisms from domestic wastewater by aerated lagoons. *Canadian Journal of Microbiology* 56(2), 188–194. [PubMed: 20237581]
- Lodder WJ and de Roda Husman AM (2005) Presence of noroviruses and other enteric viruses in sewage and surface waters in The Netherlands. *Applied and Environmental Microbiology* 71(3), 1453–1461. [PubMed: 15746348]
- Love DC and Sobsey MD (2007) Simple and rapid F+ coliphage culture, latex agglutination, and typing assay to detect and source track fecal contamination. *Applied and Environmental Microbiology* 73(13), 4110–4118. [PubMed: 17483282]
- Luther K and Fujioka R (2004) Usefulness of monitoring tropical streams for male-specific RNA coliphages. *Journal of Water and Health* 2(3), 171–181. [PubMed: 15497813]
- Mandilara GD, Smeti EM, Mavridou AT, Lambiri MP, Vatopoulos AC and Rigas FP (2006) Correlation between bacterial indicators and bacteriophages in sewage and sludge. *FEMS Microbiology Letters* 263(1), 119–126. [PubMed: 16958859]
- McMinn BR, Ashbolt NJ and Korajkic A (2017) Bacteriophages as indicators of faecal pollution and enteric virus removal. *Letters in Applied Microbiology* 65(1), 11–26. [PubMed: 28304098]
- Medema GJ, van Asperen IA and Havelaar AH (1997) Assessment of the exposure of swimmers to microbiological contaminants in fresh waters. *Water Science and Technology* 35(11), 157–163.

- Mendez J, Jofre J, Lucena F, Contreras N, Mooijman K and Araujo R (2002) Conservation of phage reference materials and water samples containing bacteriophages of enteric bacteria. *Journal of Virological Methods* 106(2), 215–224. [PubMed: 12393152]
- Meuleman AFM, van Logtestijn R, Rijs GBJ and Verhoeven JTA (2003) Water and mass budgets of a vertical-flow constructed wetland used for wastewater treatment. *Ecological Engineering* 20(1), 31–44.
- Microsoft Corporation (2013) Microsoft Excel version 2013.
- Moce-Llivina L, Lucena F and Jofre J (2005) Enteroviruses and bacteriophages in bathing waters. *Applied and Environmental Microbiology* 71(11), 6838–6844. [PubMed: 16269717]
- Montazeri N, Goettert D, Achberger EC, Johnson CN, Prinyawiwatkul W and Janes ME (2015) Pathogenic Enteric Viruses and Microbial Indicators during Secondary Treatment of Municipal Wastewater. *Applied and Environmental Microbiology* 81(18), 6436–6445. [PubMed: 26162869]
- Muniesa M, Lucena F, Blanch AR, Payan A and Jofre J (2012) Use of abundance ratios of somatic coliphages and bacteriophages of *Bacteroides thetaiotaomicron* GA17 for microbial source identification. *Water Research* 46(19), 6410–6418. [PubMed: 23039916]
- Muniesa M, Balleste E, Imamovic L, Pascual-Benito M, Toribio-Avedillo D, Lucena F., Blanch AR and Jofre J (2018) Bluephage: A rapid method for the detection of somatic coliphages used as indicators of fecal pollution in water. *Water Research* 128, 10–19. [PubMed: 29078067]
- Nappier SP, Aitken MD and Sobsey MD (2006) Male-specific coliphages as indicators of thermal inactivation of pathogens in biosolids. *Applied and Environmental Microbiology* 72(4), 2471–2475. [PubMed: 16597945]
- Nappier SP, Soller JA, and Eftim SE (2018) Potable Water Reuse: What Are the Microbiological Risks? *Current Environmental Health Reports*, 5(2), 283–292 [PubMed: 29721701]
- Ogorzaly L and Gantzer C (2006) Development of real-time RT-PCR methods for specific detection of F-specific RNA bacteriophage genogroups: application to urban raw wastewater. *Journal of Virological Methods* 138(1–2), 131–139. [PubMed: 16997389]
- Pallin R, Wyn-Jones AP, Place BM and Lightfoot NF (1997) The detection of enteroviruses in large volume concentrates of recreational waters by the polymerase chain reaction. *Journal of Virological Methods* 67(1), 57–67. [PubMed: 9274818]
- Palmateer GA, Dutka BJ, Janzen EM, Meissner SM and Sakellaris MG (1991) Coliphage and bacteriophage as indicators of recreational water quality. *Water Research* 25(3), 355–357.
- Pouillot R, Van Doren JM, Woods J, Plante D, Smith M, Goblick G, Roberts C, Locas A, Hajen W, Stobo J, White J, Holtzman J, Buenaventura E, Burkhardt W 3rd, Catford A, Edwards R, DePaola A and Calci KR (2015) Meta-analysis of the reduction of norovirus and male-specific coliphage concentrations in wastewater treatment plants. *Applied and Environmental Microbiology* 81(14), 4669–4681. [PubMed: 25934626]
- R Development Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rezaeinejad S, Vergara GG, Woo CH, Lim TT, Sobsey MD and Gin KY (2014) Surveillance of enteric viruses and coliphages in a tropical urban catchment. *Water Research* 58, 122–131. [PubMed: 24747143]
- Simmons FJ and Xagoraki I (2011) Release of infectious human enteric viruses by full-scale wastewater utilities. *Water Research* 45(12), 3590–3598. [PubMed: 21570703]
- Sinclair RG, Jones EL and Gerba CP (2009) Viruses in recreational water-borne disease outbreaks: a review. *Journal of Applied Microbiology* 107(6), 1769–1780. [PubMed: 19486213]
- Skraber S, Gassilloud B and Gantzer C (2004) Comparison of coliforms and coliphages as tools for assessment of viral contamination in river water. *Applied and Environmental Microbiology* 70(6), 3644–3649. [PubMed: 15184169]
- Sobsey MD, Battigelli D, Handzel T and Schwab K (1995) Male-specific coliphages as indicators of viral contamination of drinking water, pp. 1847–1862. In *Proceedings of the 1994 Water Quality Technology Conference American Water Works Association*, Denver, Colo.
- Soller JA, Eftim SE, Wade TJ, Ichida AM, Clancy JL, Johnson T, Schwab K, Ramirez-Toro G, Nappier SP and Ravenscroft JE (2015) Use of quantitative microbial risk assessment to improve interpretation of a recreational water epidemiological study. *Microbial Risk Analysis* 1(1), 2–11.

- Soller JA, Schoen ME, Bartrand T, Ravenscroft J and Ashbolt NJ (2010) Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Research* 44(16), 4674–4691. [PubMed: 20656314]
- Soller JA, Eftim SE and Nappier SP (2018) Direct potable reuse microbial risk assessment methodology: Sensitivity analysis and application to State log credit allocations. *Water Research* 128, 286–292. [PubMed: 29107913]
- Sprenger C, Lorenzen G, Grunert A, Ronghang M, Dizer H, Selinka HC, Girones R, Lopez-Pila JM, Mittal AK and Szewzyk R (2014) Removal of indigenous coliphages and enteric viruses during riverbank filtration from highly polluted river water in Delhi (India). *Journal of Water and Health* 12(2), 332–342. [PubMed: 24937227]
- Olivieri A, Crook J, Anderson M, Bull R, Drewes J, Haas C, Jakubowski W, McCarty P, Nelson K, Rose J, Sedlak D and Wade T (2016) Expert Panel Final Report: Evaluation of the feasibility of developing uniform water recycling criteria for direct potable reuse. Report to the Legislature California.
- Stapleton CM, Wyer MD, Kay D, Crowther J, McDonald AT, Walters M, Gawler A and Hindle T (2007) Microbial source tracking: a forensic technique for microbial source identification? *Journal of Environmental Monitoring* 9(5), 427–439. [PubMed: 17492088]
- Sunger N, Hamilton KA, Morgan PM and Haas CN (2018) Comparison of pathogen-derived ‘total risk’ with indicator-based correlations for recreational (swimming) exposure. *Environmental Science and Pollution Research* 10.1007/s11356-018-1881-x.
- Teklehaimanot GZ, Coetzee MA and Momba MN (2014) Faecal pollution loads in the wastewater effluents and receiving water bodies: a potential threat to the health of Sedibeng and Soshanguve communities, South Africa. *Environmental Science and Pollution Research International* 21(16), 9589–9603. [PubMed: 24838129]
- Till D, McBride G, Ball A, Taylor K and Pyle E (2008) Large-scale freshwater microbiological study: rationale, results and risks. *Journal of Water and Health* 6(4), 443–460. [PubMed: 18401110]
- U.S. EPA (1984) Manual of Methods for Virology, 600/4–84/013 (N14), Office of Research and Development, Cincinnati, OH.
- U.S. EPA (1986) Ambient Water Quality Criteria for Bacteria - 1986, 440/5–84-002, Office of Water, Washington DC.
- U.S. EPA (2001a) Method 1602: Male-specific (F+) and Somatic Coliphage in Water by Single Agar Layer (SAL) Procedure, 821-R-01–029, Office of Water, Washington, D.C.
- U.S. EPA (2001b) Method 1601: Detection of Male-specific (F+) and Somatic Coliphage in Water by Two-Step Enrichment Procedure, 821-R-01–030, Office of Water, Washington, DC.
- U.S. EPA (2014) Microbiological Risk Assessment (MRA) Tools, Methods, and Approaches for Water Media, 820-R-14–009, Office of Science and Technology, Washington, DC.
- U.S. EPA (2015) Review of coliphages as possible indicators of fecal contamination for ambient water quality, 820-R-15–098, Office of Science and Technology, Washington, DC.
- U.S. EPA (2017) Proceedings from the US EPA Coliphage Experts Workshop March 1–2, 2016, 822-R-17–003, Office of Science and Technology, Washington, DC.
- Viau EJ, Goodwin KD, Yamahara KM, Layton BA, Sassoubre LM, Burns SL, Tong HI, Wong SH, Lu Y and Boehm AB (2011) Bacterial pathogens in Hawaiian coastal streams--associations with fecal indicators, land cover, and water quality. *Water Research* 45(11), 3279–3290. [PubMed: 21492899]
- Wade TJ, Sams E, Brenner KP, Haugland R, Chern E, Beach M, Wymer L, Rankin CC, Love D, Li Q, Noble R and Dufour AP (2010) Rapidly measured indicators of recreational water quality and swimming-associated illness at marine beaches: a prospective cohort study. *Environmental Health* 9, 66. [PubMed: 21040526]
- Wiedenmann A, Kruger P, Dietz K, Lopez-Pila JM, Szewzyk R and Botzenhart K (2006) A randomized controlled trial assessing infectious disease risks from bathing in fresh recreational waters in relation to the concentration of *Escherichia coli*, intestinal enterococci, *Clostridium perfringens*, and somatic coliphages. *Environmental Health Perspectives* 114(2), 228–236. [PubMed: 16451859]

- Wu J, Li H and Huang X (2010) Indigenous somatic coliphage removal from a real municipal wastewater by a submerged membrane bioreactor. *Water Research* 44(6), 1853–1862. [PubMed: 20045169]
- Wymer LJ and Wade TJ (2007) The Lognormal Distribution and Use of the Geometric Mean and the Arithmetic Mean in Recreational Water Quality Measurement In: Wymer LJ, ed. *Statistical Framework for Recreational Water Quality Criteria and Monitoring*. Chichester, England: John Wiley and Sons.
- Yandell BS (1997) *Practical Data Analysis for Designed Experiments*. Chapman & Hall/CRC, Taylor and Francis Group, Boca Raton, FL.

Figure. 1.
Overview of the coliphage literature search process.

Figure 2. Observed coliphage densities in raw wastewaters and in ambient waters, boxplots by geographic region and season. (A) Raw wastewater coliphage densities by geographic region; (B) Raw wastewater coliphage densities by season; (C) Ambient coliphage densities by geographic region; (D) Ambient coliphage densities by season. NA = seasonal information not available.

Figure 3. Distribution of coliphage mean density by geographic region for MSC (left) and somatic (right) coliphage in raw wastewater after bootstrapping (Stage 1).

Figure 4. Distribution of coliphage mean density by season for MSC (left) and somatic (right) coliphage in raw wastewater after bootstrapping (Stage 1).

Figure 5. Distribution of coliphage mean density in raw wastewater after bootstrapping (Stage 2) accounting for geographic region (left) and for season (right).

Table 1.

Description of coliphage in raw wastewater datasets.

Reference	Country	Number of data points ^a		Source of Data ^b
		MSC ^c	Somatic	
		Measured	Measured	
Aw and Gin 2010	Singapore	18	18	Author
Bailey et al. 2017	USA	4	4	Author
Blanch et al. 2006 ^d	Cyprus, France, Spain, Sweden, UK	112	114	Author
Carducci and Verani 2013	Italy		52	Figure 1c
Francy et al. 2012	USA	19	19	Tables S4, S5
Haramoto et al. 2006	Japan	11		Figure 2c
Haramoto et al. 2012	Japan	10		Table 1
Locas et al. 2010	Canada	24	24	Author
Lodder and de Roda Husman 2005	The Netherlands	5	5	Table 2
Mendez et al. 2002	Spain	1	1	Figure 3
Meuleman et al. 2003	The Netherlands	4		Table 7
Montazeri et al. 2015	USA	12		Figure 4
Muniesa et al. 2012	Spain		24	Figure 1
Ogorzaly and Gantzer, 2006	France	7		Figure 3
Pouillot et al. 2015	USA & Canada	170		Author
Teklehaimanot et al. 2014	South Africa		8	Table 4
Wu and Huang 2010	China		64	Figure 2a
Total		397	333	

^a Blank spaces indicate that the specific coliphage type was not evaluated.^b Author source identifies studies where the corresponding authors were contacted and provided the raw data for this study.^c MSC = male-specific coliphage^d There was one value below detection limit of 2 log₁₀ PFU/L somatic coliphage.

Table 2.

Description of coliphage in ambient waters datasets.

Reference	Location	Number of data points ^a				LOD ^b (log ₁₀ PFU/L)	Source of Data
		MSC ^c		Somatic			
		Measured	Nondetects	Measured	Nondetects		
Astrom et al. 2013	Sweden			22			Figures 2c, 2d, 3c, 3d
Brezina and Baldini 2008	Argentina			35			Figures 2, 3
Carducci et al. 2006	Italy			24			Figure 2
Davies et al. 2003	Australia	6	35	32	10	1	Figures 1, 2
DiDonato et al. 2009	USA	21	32	55	8	0	Figures 4, 5
Gourmelon et al. 2007	France	25	3			2	Table 5
Gourmelon et al. 2010	France	14	6			2	Table 7
Haramoto et al. 2009	Japan	9					Table 1
Haramoto et al. 2012	Japan	20					Tables 1, Tables 2
Jiang and Chu 2004	USA	12	9			1.3, 1.5 ^d	Table 3
Lee et al. 1997	UK	11					Table 2
Lee et al. 2011	South Korea	5					Table 1
Medema et al. 1997	Netherlands	25	5			−1	Figure 1
Moce-Llivina et al. 2005	Spain	3	17	19	1	2	Table 3
Pallin et al. 1997	England	9	6			3.3	Tables 2
Rezaeinejad et al. 2014	Singapore	11		11			Figure 5
Skraber et al. 2004	France			170			Figure 2
Sprenger et al. 2014	India			4			Table 3
Stapleton et al. 2007	England	25	34			1	Table 1
Till et al. 2008	New Zealand	234	346			1	Figure 5
Viau et al. 2011	USA	85	3			0.3	Table S2
Total		515	496	372	19		

^aBlank spaces indicate that the specific coliphage type was not evaluated.^bLOD: limit of detection for both coliphage types^cMSC = male-specific coliphage.^dDetection limit varied by sampling location.

Table 3.

Summary statistics of coliphage densities in raw wastewater. All values are log₁₀ PFU/L.

		Total (MSC ^a Somatic)			
		Number of observations	Mean	Median	Standard Deviation
Total observations		730 (393 329)	5.73 (5.50 6.00)	5.90 (5.74 6.06)	1.21 (1.11 1.26)
Observations by geographic region	North America	276 (229 47)	5.78 (5.89 5.23)	6.15 (6.20 5.28)	1.16 (1.10 1.32)
	Europe	325 (129 196)	5.32 (4.73 5.71)	5.11 (4.72 5.73)	1.16 (0.83 1.19)
	Africa	8 (- 8)	6.14 (- 6.14)	6.25 (- 6.25)	0.72 (- 0.72)
	Asia	121 (39 82)	6.70 (5.83 7.11)	6.92 (5.81 7.37)	0.81 (0.45 0.58)
Observations by season	Spring	176 (89 87)	6.07 (5.52 6.63)	6.22 (5.81 7.25)	1.26 (1.02 1.24)
	Summer	98 (43 55)	5.87 (5.49 6.17)	6.20 (5.86 6.50)	1.29 (1.29 1.22)
	Fall	62 (42 20)	6.27 (6.13 6.57)	6.30 (6.23 6.68)	0.80 (0.88 0.51)
	Winter	158 (106 52)	6.45 (6.28 6.81)	6.47 (6.29 6.94)	0.80 (0.76 0.76)

^aMSC = male-specific coliphage.

Table 4.

Summary statistics of detected coliphage densities in ambient waters. All values are log₁₀ PFU/L.

		Total (MSC Somatic)			
		Number of detected observations	Mean	Median	Standard Deviation
Total observations		887 (515 372)	2.77 (2.20 3.56)	2.52 (2.06 3.76)	1.22 (0.91 1.16)
Observations by geographic region	North America	173 (118 55)	2.24 (2.19 2.34)	2.26 (2.24 2.29)	0.77 (0.75 0.80)
	Europe	347 (112 235)	3.53 (2.58 3.98)	3.78 (2.46 4.12)	1.28 (1.31 0.98)
	Asia	60 (45 15)	3.27 (3.16 3.59)	2.90 (2.90 2.77)	1.03 (0.70 1.68)
	South America	35 (- 35)	3.46 (- 3.46)	3.41 (- 3.41)	0.62 (- 0.62)
	Oceania	272 (240 32)	1.95 (1.85 2.66)	1.82 (1.77 2.37)	0.65 (0.52 1.04)
Observations by season	Spring	64 (52 12)	2.64 (2.21 4.49)	2.28 (2.13 4.88)	1.45 (0.95 1.81)
	Summer	93 (80 13)	2.63 (2.30 4.71)	2.26 (2.00 4.54)	1.51 (1.20 1.59)
	Fall	60 (33 27)	3.11 (3.16 3.06)	2.73 (3.00 2.47)	1.30 (1.20 1.51)
	Winter	173 (88 85)	2.62 (2.49 2.75)	2.45 (2.48 2.41)	1.14 (0.91 1.34)

^aMSC = male-specific coliphage.

Table 5.

Summary statistics for the bootstrapped coliphage densities in raw wastewater.

		MSC Somatic		
Analysis		N	Mean (95% CI)	Standard Deviation (95% CI)
Stage 1				
Geographic region	North America	229 47	5.9 (5.7, 6.0) 5.2 (4.9, 5.6) **	1.1 (1.0, 1.2) 1.3 (1.1, 1.5)
	Europe	129 196	4.7(4.6, 4.9) 5.7 (5.5, 5.9) **	0.8 (0.7, 1.0) 1.2 (1.1, 1.3)
	Africa	- 8	- 6.1 (5.7, 6.6)	- 0.7 (0.6, 1.0)
	Asia	39 82	5.8 (5.7, 6.0) 7.1 (7.0, 7.2) **	0.4 (0.3, 0.6) 0.6 (0.5, 0.7)
Season	Spring	89 87	5.5 (5.3, 5.7) 6.6 (6.4, 6.9) **	1.0 (0.8, 1.3) 1.2 (1.0, 1.5)
	Summer	43 55	5.5 (5.1, 5.9) 6.2 (5.9, 6.5) *	1.3 (1.0, 1.6) 1.2 (1.0, 1.5)
	Fall	42 20	6.1 (5.9, 6.4) 6.6 (6.3, 6.8)	0.9 (0.6, 1.2) 0.5 (0.4, 0.7)
	Winter	106 52	6.3 (6.1, 6.4) 6.8 (6.6, 7.0) **	0.8 (0.6, 1.0) 0.8 (0.7, 0.9)
Stage 2				
	Accounting for geographic region		5.8 (5.4, 6.1) 6.0 (5.7, 6.4)	1.0 (0.7, 1.4) 1.3 (1.0, 1.6)
	Accounting for season		5.9 (5.5, 6.1) 6.5 (6.2, 6.8)	1.0 (0.7, 1.4) 1.0 (0.7, 1.3)

^aSignificant comparisons between MSC and somatic means by regions and seasons using t-tests ($\alpha=0.05$):*
p-value < 0.05**
p-value < 0.001. All values are log₁₀ PFU/L.

MSC = male-specific coliphage.

Table 6.

Estimated parameters for the distribution of coliphage densities in ambient water using three statistical methods.

Method	Coliphage type	N	N<LOD	Distribution Parameters	
				Mean (95% CI)	Standard Deviation
Detected data only	MSC	515	0	2.2 (2.1, 2.3)	0.9
	Somatic	372	0	3.6 (3.4, 3.7)	1.2
Kaplan –Meier	MSC	1,011	496	1.6 (1.5, 1.7)	1.8
	Somatic	391	19	3.4 (3.3, 3.6)	1.3
Maximum Likelihood (Lognormal)	MSC	1,011	496	1.4 (1.3, 1.5)	1.1
	Somatic	391	19	3.4 (3.3, 3.5)	1.3

LOD: limit of detection. All values are log₁₀ PFU/L. MSC = male-specific coliphage.