Environmental Science
Processes & Impacts

Risk-based water quality thresholds for coliphage in surface waters: Effect of temperature and contamination aging

Journal:	*Environmental Science: Processes & Impacts*
Manuscript ID	EM-ART-08-2019-000376.R1
Article Type:	Paper
Date Submitted by the Author:	19-Sep-2019
Complete List of Authors:	Boehm, Alexandria; Stanford University, Civil and Environmental Engineering:

Microbial pollution in natural waters can cause illness in swimmers and others who come into contact with the water. Coliphages, viruses that infect *Escherichia coli*, have been used for decades to assess surface water quality, but there are no clear guidelines as to what their acceptable concentrations should be in order to ensure that waters are of good enough quality for swimming. The study uses a risk-based framework to gain insight into the risk of illness associated with recreational exposure to coliphages from sewage in surface waters. We specifically explore how aging of contamination and temperature of the water affect simulated risk of illness and associated concentrations of coliphage by considering first-order decay of pathogens and coliphages in the model.

Risk-based water quality thresholds for coliphage in surface waters: Effect of temperature and contamination aging

Alexandria B. Boehm

Department of Civil & Environmental Engineering, Stanford University, Stanford, California, USA 94305-4020

A manuscript for

Environmental Science: Processes and Impacts

Abstract

Coliphages, viruses that infect *Escherichia coli*, have been used for decades to assess surface water quality yet there is no guideline for interpreting their concentrations. The present study uses a quantitative microbial risk assessment (QMRA) framework to derive risk-based surface water quality thresholds for somatic and F+ or male-specific coliphage. The risk-based threshold is the concentration at which the risk of gastro-intestinal illness is simulated to be 32/1000. The framework specifically investigates a simplified hazard scenario where recreational swimmers come into contact with water contaminated with untreated sewage containing coliphages and enteric pathogens. The framework considers exposure to sewage of diverse ages and thus accounts for the decay of coliphages and pathogens over time. As decay rate constants depend on temperature, the model considers the effect of temperature on the risk-based threshold. When exposure to fresh, unaged sewage contamination occurs, the risk-based water quality threshold for somatic and F+ coliphage is 60 PFU/100 ml and 30 PFU/100 mL, respectively, and temperature independent. The risk-based threshold decreases as the contamination ages because, on average, coliphage decay more quickly than norovirus, the pathogen that contributes the most to risk. The decrease in the risk-based threshold with contaminant age is equal to the difference in the first order decay rate constants of coliphage and norovirus. Since coliphage decay rate constants are larger at 25°C than at 15°C, and norovirus decay rate constants are a weak function of temperature, risk-based thresholds decrease more quickly with age at 25°C than at 15°C. For the common case where the age of contamination is unknown, the risk-based threshold for both coliphage is between ~1 PFU/100 ml and ~10 PFU/100 mL, depending on model assumptions. Future work can apply this QMRA framework for identifying risk-based thresholds for coliphage

from different hazards (treated wastewater or animal feces) or from mixtures of contamination of different ages and sources.

Introduction

Fecal indicator bacteria (FIB), specifically *Escherichia coli* and enterococci, are used to monitor surface water quality globally. FIB concentrations in surface waters correlate to risk of gastrointestinal illness (GI) of swimmers when FIB come from sewage, as evidenced by a wide-range of epidemiology studies[1]. They are also used to assess coastal water quality within shell fisheries[2,3], and to assess water quality of surface waters used as drinking water sources [4].

Besides sewage, FIB can come from a variety of other sources including animal feces, and environmental reservoirs, such as sand and decaying vegetation[5]. When FIB come from sources other than human sewage, their concentration in surface waters may not indicate the same level of risk to a swimmer as an equivalent FIB concentration from a different source [6]. Alternative indicators have been sought for decades to overcome the lack of FIB source specificity. Coliphages including somatic and male-specific or F+ coliphage represent examples of such indicators. Somatic coliphage are bacteriophage that infect *E. coli* through their outer membrane and F+ (or equivalently male-specific or F-specific) coliphage infect *E. coli* via pili appendages. Since they are viruses, they likely better mimic the environmental fate and transport of human viruses, considered the most important etiology of waterborne illness [7–10], than bacteria. USEPA may develop coliphage ambient water quality criteria for surface waters[10], but no criterion has been published yet.

The goal of the present study is to derive risk-based water quality thresholds for coliphage in surface waters using a modeling approach. Quantitative microbial risk assessment (QMRA) has previously be used to derive "risk-based thresholds" for alternative fecal indicators in surface waters including HF183, HumM2, and crAssphage, human feces-associated DNA markers [11–14] and a gull feces-associated DNA marker[15,16]. The approach uses a risk-framework

to estimate the concentration of the fecal indicator in surface waters that would result in the probability of illness (approximately 32/1000 swimmers) used by USEPA to guide development of surface water FIB criteria[17].

In the first iterations of our previous work, risk-based thresholds were derived for HF183, HumM2, and a gull-feces associated marker assuming that the indicator was from sewage or feces and that the contamination was "fresh" and not aged[12,16]. In the second iteration of that work, we explored how the risk-based threshold for HF183 changes as the contamination ages, and determined what the risk-based threshold would be if the age of contamination was unknown (which is typically the case)[11]. We found that the risk-based threshold decreases as the contamination ages. The second iteration required information on the decay rate constants for all reference pathogens used in the QMRA as well as for HF183. Based on a systematic review and meta-analysis of those decay rate constants, we constructed distributions of the rate constants for use in the QMRA Monte-Carlo simulations. Although the meta-analysis indicated that the rate constants were a function of temperature, we did not consider their temperature dependence in that analysis.

The present study extends our previous work in several ways. First, we explicitly consider the effect of temperature on the QMRA-derived risk-based thresholds. Second, we extend the approach to deduce the risk-based threshold for somatic and male-specific F+ coliphage from sewage of different ages. The extension of the approach to coliphage is possible given a recent systematic review and meta-analysis of mammalian virus and coliphage decay rate constants in surface waters[18] as well as a recent systematic review of their concentrations in untreated sewage[19]. The work described herein is guided by the following research question: What is the risk-based threshold for coliphages in cold versus warm water? We assume that the

source of the coliphages is raw sewage, and we consider the possibility that the contamination is aged a known amount of time, and the case where it is potentially aged, but the age is unknown.

Methods

QMRA: Exposure to untreated, raw sewage of known age. A static QMRA was used to estimate gastrointestinal illness risk from swimming in surface waters with varying concentrations of coliphage from untreated sewage of different ages using R (Version 1.1.463).. The influence of immunity and secondary transmission was not considered in the models[20]. Swimmer exposure to microbially-contaminated waters may also lead to other symptoms including respiratory illness and skin rash[5,21] that are not considered in the model. In the QMRA, somatic coliphage or F+ coliphage serves as an index for the amount of sewage present in surface water. Methods generally mirror those used by Boehm et al.[12] who investigated risk-based thresholds for HF183 for recreational exposure to sewage of diverse ages. The QMRA considers the cumulative risk from exposure to reference pathogens adenovirus (not considered by Boehm et al. [12]), norovirus, *Giardia, Cryptosporidium, E. coli* O157:H7, *Campylobacter*, and *Salmonella* as recommended by USEPA and used extensively in bathing water QMRAs[6,22–24]. The technique uses 10,000 Monte Carlo simulations to randomly draw model parameters from their respective distributions for each model scenario.

The indicator (somatic coliphage or F+ coliphage) concentration (C_{meas}) in surface water serves as an input to the model and varies between 10^{-3} and 10^{3} PFU/100 ml in 0.2 order of magnitude increments. The upper end of the range was chosen to be less than the concentrations of the indicators in sewage (Table 1). The lower ends were chosen to be at or below a typical lower measurement limit given common methods used for the indicator detection. The age of the

contamination (the time it has spent in surface water after being released from an untreated, raw sewage source) is τ. τ serves as a model input and varies between 0 (unaged) and 15 days (d) in 0.5 d increments. A total of 961 C_{meas}-τ combinations (31 distinct ages and 31 distinct indicator concentrations) were modeled for each of the 2 coliphages at the 2 temperatures for a total of 3844 combinations of C_{meas}-τ.

After C_{meas} and τ are specified as model inputs, the concentration of i^{th} reference pathogen in surface waters ($C_{i\_surface}$) is modeled as follows:

$$C_{i\_surface} = \frac{C_{meas} C_{i\_sewage}}{C_{indicator\_sewage}} e^{\Delta k \tau} \quad (1)$$

where $\Delta k = k_{indicator} - k_i$, and C_{i_sewage} and $C_{indicator\_sewage}$ are, respectively, the i^{th} reference pathogen and indicator concentrations in sewage, and k_i and $k_{indicator}$ are their first order decay rate constants. In words, equation 1 illustrates that in order to infer the concentration of pathogen *i* in surface waters given the measured concentration of an indicator, one must know the relative decay rate constants of the pathogen and indicator, the age of contamination, and the concentration of indicator and pathogen in the contamination source (in this case sewage). In equation 1, C_{i_sewage} and $C_{indicator\_sewage}$ are described by distributions (Table 1), assumed to be independent, that were obtained from the literature and have been used in previous QMRA studies[15,25].

First order decay rate constants of allochthonous microorganisms in surface waters are generally a function of water temperature[11,18]. We considered two temperatures, 15°C and 25°C. These two temperatures represent water temperatures typical of a cooler and warmer recreational water like those in central California, and Hawai'i, respectively. Previous work compiled surface water decay rate constants of the indicators and pathogens used in this study for surface as well as the experimental conditions, including temperatures, at which the rate constants were

derived[11,18]. We derived distributions for the rate constants needed in equation 1 ($k_{indicator}$ and k_i) for temperatures of 15°C and 25°C using the methods outlined in the next section and then $k_{indicator}$ and k_i values in surface water were randomly drawn from their respective distributions. In deriving equation 1, it is assumed that the advection and dispersion of indicator and reference pathogens are identical, and any non-conservative behavior of targets is adequately captured by first order kinetics.

During each model run, it was ensured that $F = C_{meas}\exp(k_{indicator}\tau)/C_{indicator\_sewage}$ (the volume fraction of sewage present in the water) did not exceed 1 and if it did, then new model parameters were drawn from their respective distributions [15]. If $F \geq 1$ for more than 10% of the Monte Carlo draws for a particular C_{meas}-τ combination, the combination was deemed unlikely. Practically, this implies that if the source of the indicator is sewage, and the indicator decays over time, there is a time at which a measurement of a relatively high concentration of indicator is unlikely. The relationship between C_{meas} and τ defining this boundary is presented.

It is assumed that the volume (V) of water ingested by a swimmer per swimming event follows the $\log_{10}$-normal distribution with a mean of 1.20 and standard deviation of 0.68; units of V are ml [26]. The mean and standard deviation were derived using the percentile values reported by Deflorio-Barker et al. [26] for all swimmers (including children), and assuming the data are log-normally distributed to relate the percentiles to the standard deviation using a table of Z values. The dose of pathogen i, μ_i, is given by $C_{i\_surface}V$. The dose was used as input to the reference pathogen dose-response functions (Table 1) to determine the probability of infection (P_{inf_i}). The probability of illness (P_{ill_i}) was calculated by multiplying the probability of infection by the probability of illness given infection $P_{ill|inf\_i}$ (Table 1). $P_{ill|inf\_i}$ was randomly drawn from a uniform distribution for each model iteration except for the case of *Campylobacter* which used a

dose-dependent formula (Table 1). The cumulative risk of illness from exposure to all reference pathogens (P_{ill}) is given by $P_{ill} = 1 - \prod_i (1 - P_{ill\_i})$. It was assumed that infection and illness for each pathogen is independent.

10,000 iterations were obtained for each C_{meas}-τ combination. For some C_{meas}-τ combinations, more than 10,000 draws were needed as in some cases, F was greater than 1 and the model parameters were redrawn, but the total number of draws was not allowed to exceed 15,000. The median, interquartile range, and 10th and 90th percentiles of P_{ill} for each C_{meas}-τ combination were calculated from the respective 10,000 iterations. P_{ill} was compared to the value 32/1000 which is equal to the risk threshold published by USEPA for bathing water for a single swimming event[17].

Temperature-specific rate constant distributions. Compiled *k* values for the reference pathogens and indicators were obtained from two previous systemic reviews[11,12] along with the temperatures at which the experiments were conducted and the analytical method used to enumerate the target organisms (these data are all available as supplemental information in the cited papers). Norovirus decay rate constants were reported in both papers; the compilation from Boehm et al. [18] are used herein. Norovirus decay rate constants are estimated from all data available for viruses in the species *Norwalk virus*, which includes murine and human norovirus [18]. In this work, we only considered the temperature dependence of *k*, and not dependences on other experimental factors (such as sunlight and salinity) for the purpose of characterizing the plausible values of *k*. We first modeled $\log_{10}$-*k* as a function of temperature (continuous variable) and analytical enumeration method (categorical variable) using the linear model or (lm)function in R. If the analytical enumeration method was not statistically significant in the model ($p>0.05$),

then a reduced model was created that only considered temperature as an independent variable. Using the linear model coefficients and intercepts, we used the 'predict.lm' to predict the estimated mean and standard deviation of $\log_{10}$-k for the indicators and all reference pathogens for temperatures of 15°C and 25°C as measured using culture methods. Note that since norovirus k values were modeled with all *Norwalk virus* species k values including those of cultivatable murine norovirus[18], an estimate for norovirus k measured by culture methods is possible. If k was not a function of temperature according to the linear model ($p>0.05$), then the estimated mean and standard deviation of the entire compiled dataset of k values was used to represent k at both temperatures.

QMRA: Exposure to untreated sewage of unknown age. The age of sewage in the surface water is usually unknown and furthermore, surface water contamination may represent a mixture of contamination of diverse ages. We repeated the QMRA for a scenario where the indicator is measured in surface waters but the contamination age is unknown. C_{meas} was specified as a model input at the same values used above. τ was drawn from a uniform distribution ranging from 0 to a maximum realistic value (τ_{max}) given the specified C_{meas} (derived in results section and related to the F described previously). An alternative relationship between τ_{max} and C_{meas} that relates the two through the median concentration of coliphage in raw sewage and their median decay rate constant is also used to investigate the sensitivity of the result to this relationship. All other QMRA methods were the same as those above. 10,000 iterations were run for each C_{meas} to obtain distributions of P_{ill}.

Effect of decay rate constant on risk-based threshold. We investigated how the decay rate constant of an arbitrary indicator ($k_{indicator}$) affects the risk-based threshold of the indicator as a function of τ using a simplified QMRA. The QMRA followed the methods outlined above except $k_{indicator}$ was varied as a constant in $\log_{10}$ increments of 0.05 from a minimum of $\log_{10}k_{indicator} = -1.2$ to a maximum of 0.3 where $k_{indicator}$ has units d^{-1}. For each value of $k_{indicator}$, simulations were run for τ = 0 to 7 d in 0.5 day increments, and C_{meas} (the concentration of the indicator) was varied between 10^{-3} and 10^{3} per 100 ml in 0.5 $\log_{10}$ unit increments. We used $C_{indicator\_sewage}$ of F+ coliphage for the simulations. 1000 iterations were run for each combination of $k_{indicator}$, C_{meas}, and τ. The slope of the line fit to natural log (ln)-transformed risk-based threshold versus τ and median Δk for each $k_{indicator}$ value were compiled to investigate the relationship between those parameters.

Results

$\log_{10}k$ distributions. $\log_{10}$-transformed surface water k ($\log_{10}k$) values for norovirus, adenovirus, *Salmonella*, *Campylobacter*, *E. coli* O157:H7, *Giardia*, *Cryptosporidium*, somatic coliphage and F+ coliphage were modeled as a function of temperature and analytical enumeration method using the 'lm' function in R. Analytical enumeration method was not a statistically significant variable in the models for norovirus, adenovirus, *Salmonella*, or *Cryptosporidium*, so reduced models only considered the effect of temperature. Decay experiments for *E. coli* O157:H7, somatic coliphage, and *Giardia* were all completed using a single analytical method (culture methods), so temperature alone was considered as an independent variable in those models. Temperature was a significant variable in all the $\log_{10}k$

models. Using 'predict.lm' we predicted the estimated mean and standard deviation of $\log_{10}k$ values for the different organisms at temperature of 15°C and 25°C (Table 2).

τ_{max} and $C_{meas|max}$. Assuming the indicator source is raw sewage, there is an upper limit on the age of the indicator that is a function of its measured concentration in the environment C_{meas}. This may also be conceptualized as the maximum possible measured concentration $C_{meas|max}$ given τ. τ_{max} was determined for each tested C_{meas} as the maximum tested τ at which $F < 1$ in 90% or more of the simulations. A linear regression between $\log_{10}C_{meas|max}$ and τ_{max} yielded a best fit equation with slope m_1 and intercept b_1 describing their relationship ($R^2>0.99$ for each indicator at each tested temperature) (Table 3). The linear equation can be rearranged to yield an empirical relationship between τ_{max} and $\log_{10}C_{meas|max}$:

$$\tau_{max} = 1/m_1 \, [\log_{10}C_{meas|max} - b_1] \quad (2)$$

An alternative approach to estimating τ_{max} given C_{meas} is to consider conceptually a limiting case where a surface water is 100% raw sewage. At $\tau=0$, $C_{meas|max}$ can be best approximated as the median indicator concentration in sewage (Table 1). As time progresses, $C_{meas|max}$ decreases exponentially according with a first order rate constant best approximated by the indicator's median k value. In this scenario,

$$\tau_{max} = 1/m_2 \, [\log_{10}C_{meas|max} - b_2] \quad (3)$$

where b_2 is the $\log_{10}$-transformed median concentration of the indicator in raw sewage, and m_2 is $-k_{med}*\log_{10}e$ where k_{med} is the median of the indicator decay rate constant in surface waters (Table 2).

Simulated illness as a function of coliphage concentrations and τ. P_{ill} increases as the measured concentration of coliphage of a fixed age increases in ambient waters (Figure 1). Illness from exposure to norovirus contributes the most to P_{ill} (Figure 2 illustrates this for somatic coliphage at 15°C, other indicator-temperature scenarios are not shown, but results are the same). The risk-based threshold for both somatic and F+ coliphage decrease as a function of τ for T=15°C and 25°C (Figure 3), due to the fact that coliphage *k* is higher, on average, than norovirus *k*.

At T = 15°C, the risk-based threshold for somatic coliphage decreases from 60 PFU/100 ml at τ=0 to less than 0.1 PFU/100 ml at τ ~ 14 days, and continues to decrease to 0.06 PFU/100 ml at τ=15 d. The risk-based threshold for F+ coliphage decreases from 30 PFU/100 ml at τ=0 to less than 0.1 PFU/100 ml at τ~12 days; it continues to decrease to 0.02 PFU/100 mL at τ =15 d.

The decrease in coliphage risk-based threshold values with τ is greater at T = 25 °C compared to T = 15°C. At 25°C for somatic coliphage, between 0 and 4 d, the thresholds decrease from 60 to 0.03 PFU/100 ml. For F+ coliphage, the thresholds decrease from 30 to 0.0006 PFU/100 ml between 0 and 6 d. Threshold values are not reported for somatic and F+ coliphages beyond the times mentioned in this paragraph as the threshold values become greater than $C_{meas|max}$ meaning they are unrealistic.

The differences in the trends between the risk-based thresholds and τ at the two temperatures are due to the differing effects of temperature on the coliphages and norovirus *k*, norovirus *k* has a weaker temperature dependence than both coliphage *k* (Table 2). For somatic coliphage, median Δk ($k_{somatic\_coliphage}$- k_{noro}) across all simulations is 0.5 d^{-1} at 15°C and 1.6 d^{-1} at 25°C. For F+ coliphage, median Δk ($k_{F+coliphage}$- k_{noro}) across all simulations is 0.5 d^{-1} at 15°C and

1.8 d^{-1} at 25°C. The slope of the line describing the relationship between ln-transform risk-based thresholds and τ nearly match those of Δk (Figure 4).

To further explore the relationship between Δk and the slope describing the change in the risk-based threshold with τ, we carried out the QMRA simulations while varying $k_{indicator}$ systematically as a constant. The results of those simulations support the finding that the change in the ln-transformed risk-based threshold with the age of contamination is equal to Δk (Figure 4).

Risk-based threshold of coliphage for contamination of unknown age. For most contamination scenarios, the age of the contamination is unknown. In fact, contamination likely represents a mixture of contamination of diverse ages, so assignment of a single age may be unrealistic. We repeated the QMRA but instead of specifying τ, we defined τ as a random variable described by a uniform distribution bounded by 0 as a minimum value and τ_{max} as a maximum value. Recall τ_{max} represents the maximum possible age of contamination given C_{meas} as described by equation 2 or 3.

Median P_{ill} as a function of C_{meas} the two coliphages when the age of contamination is unknown is shown in Figure 5. Log_{10}-transformed median P_{ill} is approximately linear with $\log_{10}C_{meas}$ until C_{meas} is relatively high. We therefore fit a line to the region that was linear and which clearly encompassed the point where the 32/1000 risk line intercepts the curve, and then used that best fit line to calculate risk-based thresholds for the indicators when the age of contamination is unknown. Assuming equation 2 describes the relationship between $C_{meas|max}$ and τ_{max}, risk-based threshold is 14 PFU/100 ml for somatic coliphage and 3 PFU/100 ml for F+ coliphage regardless of temperature. If equation 3 is instead used to describe the relationship

between $C_{meas|max}$ and τ_{max}, then the risk-based threshold is 1 PFU/100 ml for somatic coliphage and 0.5 PFU/100 ml for F+ coliphage regardless of temperature.

Discussion

This study presents risk-based water quality thresholds for coliphage in surface waters assuming the source of contamination is raw sewage. When contamination is unaged and unaffected by first-order decay processes, then the risk-based thresholds for somatic and F+ coliphage are 60 and 30 PFU/100 ml, respectively. The threshold for somatic coliphage is higher as its concentration in sewage tends to be higher than that of F+ coliphage in sewage[19]. The threshold decreases with the age of contamination exponentially; that is, there is a linear relationship between ln-transformed threshold and τ. The slope of that line is equal to the median $k_{indicator} - k_{noro}$. This is because norovirus contributes the most to the total risk in the simulations. If other dose-response curves are used for the QMRA pathogens or different distributions are used for the concentrations of the pathogens in sewage, then this relationship may not hold.

After a certain amount of aging in surface waters, the thresholds decrease to concentrations that are lower than most standard assay detection limits. For example, if one liter of water is assayed for coliphage, the lowest detectable concentration is 0.1 PFU/100 mL (or 1 PFU/L). Depending on the water temperature and the type of coliphage, the risk-based threshold decreased below 0.1 PFU/100 ml when the contamination is aged between 3.5 and 14 days old. For the common scenario where the age of contamination is unknown, the risk-based threshold for both coliphage is also relatively low – between ~1 and 10 PFU/100 mL at both temperatures. Together, these results suggest that when raw sewage is the source of contamination, even low

concentrations of coliphage near the detection limit of standard assays may indicate a meaningful health risk (near 32/1000).

These QMRA-derived thresholds are consistent with results of several epidemiology studies that investigated the relationship between swimmer health and coliphage concentration . Colford et al.[26] and Wade et al.[28] showed detection of F+ coliphage at low levels may be associated with increased gastrointestinal illness in swimmers, consistent with the results presented herein. Wiedenmann et al.[27] found "no observed adverse effects" on swimmer health when somatic coliphages were at concentrations less than ~10 PFU/100 mL at a freshwater beach in Germany, which is close to the simulated risk-based threshold in this study when sewage is unaged. A statistically unsubstantiated relationship with low levels of somatic coliphage and swimmer health was presented by Abdelzaher et al.[29]

Overall, the risk-based thresholds are lower than those previously identified for HF183 and HumM2, which are human-associated DNA markers of fecal contamination used for assessing surface water quality. Those thresholds were on the order of 10^3 to 10^4 gene copies per 100 ml[13]. The differences are mostly attributable to the fact that concentrations of the human markers in sewage are higher than concentrations of coliphage in sewage[19,28,29].

Temperature had a noticeable effect on the coliphage risk-based thresholds with the thresholds being lower, for a specific age, at 25°C compared to 15°C, sometimes orders of magnitude lower. This is because the difference in first order decay rate constants between norovirus and coliphage are greater for waters at 25°C than 15°C. The small differences in their decay rates at 15°C results in the risk-based threshold decreasing slightly with τ (0.5 ln units per day) whereas the decrease is ~1.7 ln units per day at 25°C. These changes in the risk-based threshold with τ are equal to median Δk across the simulations used to deduce the thresholds.

When the age of contamination is unknown, however, there is no difference between the risk-based thresholds of the coliphage across temperatures. This is because when contamination is unknown, the model considers all possible ages given the measured concentration of coliphage in surface water. This added variability in the model output is greater than that introduced by considering the effect of temperature on *k*. To illustrate the sensitivity of the risk-based threshold for unknown aged contamination to the relationship between $C_{meas|max}$ and τ_{max}, we considered two possible relationships between $C_{meas|max}$ and τ_{max}. One relationship is relatively conservative in that it assumes relatively low concentrations of the indicator in sewage and relatively large decay rate constants, while the other uses best estimates of the two variables by using their medians. The risk-based threshold when contamination age is unknown changes by an order of magnitude depending on which relationship is used.

The "unknown age" scenario considered here calculates the risk of illness for all possible contamination ages given C_{meas} and then identifies their median and uses that to identify the risk-based threshold. As such, it might also conceptually represent a scenario where there is a mixture of contamination with equal proportions of all possible ages. Future work should consider the risk-based threshold for mixtures of contamination of diverse ages, but not in equal proportions to determine if there is some limiting case that may be useful for identifying an actionable risk-based water quality threshold for management of surface waters. An additional consideration is that the maximum age of contamination may be constrained by the residence time of water at a particular recreational site. For example, the residence time of water in a small bay might be at most three days. In this case, a simulation could be run where τ_{max} is not allowed to exceed 3 days to calculate a more refined estimate of the risk-based threshold when the age of contamination is unknown.

What does it mean when the risk-based threshold for a contamination of a particular age is higher than $C_{meas|max}$? Figure 3 shows that this occurs for T=25°C when $C_{meas|max}$ goes as τ according to equation 2. In a previous study[11], we observed that this occurred for the HF183 human-associated fecal marker risk-based threshold after τ~3 days, and we interpreted it to mean that there was not a risk of illness greater than 30/1000 after ~3 days. After some reflection, this interpretation is flawed. Rather, when the risk-based threshold of an indicator is higher than its $C_{meas|max}$, it means that the indicator decays more quickly than the risk of illness as the contamination ages. If this occurs too soon (when τ is small and less than the residence time of water in a recreational water body, for example), then the indicator may not be suitable for assessing risk. A challenge is determining the best approximation for representing the relationship between $C_{meas|max}$ and τ_{max}.

The results of this study provide a number of insights into the attributes of a good indicator for gastrointestinal illness risk in recreational waters. Some of these insights are not novel, as the perfect indicator for fecal contamination has been described in the literature for decades[30]. First, the indicator needs to be in sufficiently high concentrations in sewage such that once released into surface waters it persists at detectable levels for days to weeks. Second, ideally the indicator would have the same decay rate constant as the pathogen that is contributing the most to risk. In this case, Δk would be equal to 0, and the risk-based threshold would be the same regardless of contamination age, eliminating uncertainties associated with choosing a threshold for use as a risk-management tool. An obvious indicator would be norovirus itself. Third, assuming the indicator is not norovirus, k of the indicator should not be much larger than Δk or the risk-based threshold for the indicator will be larger than the maximum expected

concentration of the indicator in surface water at small ages. Conceptually, k and Δk control where the lines in Figure 3 panels intersect.

We conducted an extensive sensitivity analysis for a nearly identical QMRA model in a previous publication[11] so such an analysis is not presented herein. The QMRA model was reported to be most sensitivity to C_{noro_sewage}, $C_{indicator\_sewage}$, k_{noro}, V, $k_{indicator}$, and τ. It is important to note that there are other published dose-response curves for norovirus[31], *Campylobacter*[32], *Salmonella*[33], and *Cryptosporidium*[34] that we did not consider in the present paper. The choice of dose-response curves used in the present study mirror those used in projects that harmonize QMRA predictions with swimmer epidemiology study findings[22,35]. However, as exposure to norovirus contributed the most to total risk in this study, we repeated the QMRA replacing the norovirus dose-response model in Table 1 with a weighted version of the two available norovirus dose-response curves following the approach used by others[15,22] (see supporting information). Using the weighted norovirus dose-response functions gives slightly higher risk-based thresholds for known τ (2 times higher) and unknown τ (2-6 times higher) (Figure S1 and Tables S1 and S2). While exposure to norovirus commonly dominates QMRAs in recreational water, estimates of infectious norovirus concentrations and decay in environmental waters are highly uncertain due to the lack of a human norovirus culture system that can be used for testing environmental media. A recent study showed swimmers exposed to marine recreational waters were infected with norovirus supporting the importance of this recreational exposure for norovirus transmission[36]. Norovirus is self-limiting and not a reportable illness in the United States, so it is difficult to assess its contribution to waterborne illness using public health data. The QMRA considered pathogens using an approach that has been applied successfully in other bathing water risk studies. However, there are other pathogens that may

contribute to risk that were not considered including enteroviruses and *Shigella*. The QMRA used the best available information at the time of model implementation, but can be updated to reflect new findings on pathogen and indicator distributions, dose-response curves, and exposure assessments. This QMRA considered a specific hazard, water contaminated with untreated sewage, and thus the results should be cautiously extended to other hazards such as swimming in water contaminated by coliphage from a mixture of diverse sources such as treated wastewater effluent and animal feces.

Acknowledgements. AB was partially supported by the National Science Foundation (CBET-1804169). Jeff Soller, Mary Schoen, and Sharon Nappier provided valuable feedback on some aspects of this work.

(1) Boehm, A. B.; Soller, J. A. Recreational Water Risk: Pathogens and Fecal Indicators. In *Environmental Toxicology: Selected Entries from the Encyclopedia of Sustainability Science and Technology*; Laws, E. A., Ed.; Springer New York: New York, NY, 2013; pp 441–459.

(2) Vaughn, J. M.; Metcalf, T. G. Coliphages as Indicators of Enteric Viruses in Shellfish and Shellfish Raising Estuarine Waters. *Water Res.* **1975**, *9* (7), 613–616. https://doi.org/10.1016/0043-1354(75)90165-7.

(3) NOAA. *Classified Shellfish Growing Waters*; NOAA's State of the Coast Report. Silver Spring; MD, 1998.

(4) WHO. Drinking-water Fact Sheet http://www.who.int/mediacentre/factsheets/fs391/en/ (accessed Sep 28, 2017).

(5) Boehm, A. B.; Ashbolt, N. J.; Colford Jr., J. M.; Dunbar, L. E.; Fleming, L. E.; Gold, M.; Hansel, J.; Hunter, P. R.; Ichida, A. M.; McGee, C.; et al. A Sea Change Ahead for Recreational Water Quality Criteria. *J. Water Health* **2009**, *7*, 9–20.

(6) Soller, J. A.; Schoen, M. E.; Bartrand, T.; Ravenscroft, J. E.; Ashbolt, N. J. Estimated Human Health Risks from Exposure to Recreational Waters Impacted by Human and Non-Human Sources of Faecal Contamination. *Water Res.* **2010**, *44* (16), 4674–4691.

(7) Kotloff, K. L.; Nataro, J. P.; Blackwelder, W. C.; Nasrin, D.; Farag, T. H.; Panchalingam, S.; Wu, Y.; Sow, S. O.; Sur, D.; Breiman, R. F.; et al. Burden and Aetiology of Diarrhoeal Disease in Infants and Young Children in Developing Countries (the Global Enteric Multicenter Study, GEMS): A Prospective, Case-Control Study. *The Lancet* **2013**, *382* (9888), 209–222. https://doi.org/10.1016/S0140-6736(13)60844-2.

(8) Scallan, E.; Hoekstra, R.; Angulo, F.; Tauxe, R.; Widdowson, M.; Roy, S.; Jones, J.; Griffin, P. Foodborne Illness Acquired in the United States--Major Pathogens. *Emerg. Infect. Dis.* **2011**, *17* (1), 7–15.

(9) Sinclair, R. G.; Jones, E. L.; Gerba, C. P. Viruses in Recreational Water-Borne Disease Outbreaks: A Review. *J. Appl. Microbiol.* **2009**, *107* (6), 1769–1780. https://doi.org/10.1111/j.1365-2672.2009.04367.x.

(10) United States Environmental Protection Agency. *Review of Coliphages as Possible Indicators of Fecal Contamination for Ambient Water Quality*; 820-R-15–098; US EPA: Washington D. C., 2015.

(11) Boehm, A. B.; Graham, K. E.; Jennings, W. C. Can We Swim yet? Systematic Review, Meta-Analysis, and Risk Assessment of Aging Sewage in Surface Waters. *Environ. Sci. Technol.* **2018**, *52* (17), 9634–9645.

(12) Boehm, A. B.; Soller, J. A.; Shanks, O. C. Human-Associated Fecal QPCR Measurements and Simulated Risk of Gastrointestinal Illness in Recreational Waters Contaminated with Raw Sewage. *Environ. Sci. Technol. Lett.* **2015**, *2*, 270–275. https://doi.org/10.1021/acs.estlett.5b00219.

(13) Zhang, Q.; Gallard, J.; Wu, B.; Harwood, V. J.; Sadowsky, M. J.; Hamilton, K. A.; Ahmed, W. Synergy between Quantitative Microbial Source Tracking (QMST) and Quantitative Microbial Risk Assessment (QMRA): A Review and Prospectus. *Environ. Int.* **2019**, *130*, 104703. https://doi.org/10.1016/j.envint.2019.03.051.

(14) Crank, K.; Petersen, S.; Bibby, K. Quantitative Microbial Risk Assessment of Swimming in Sewage Impacted Waters Using CrAssphage and Pepper Mild Mottle Virus in a

Customizable Model. *Environ. Sci. Technol. Lett.* **2019**. https://doi.org/10.1021/acs.estlett.9b00468.

(15) Brown, K. I.; Graham, K. E.; Soller, J. A.; Boehm, A. B. Estimating the Probability of Illness Due to Swimming in Recreational Water with a Mixture of Human- and Gull-Associated Microbial Source Tracking Markers. *Environ. Sci. Process. Impacts* **2017**, *19* (12), 1528–1541. https://doi.org/10.1039/C7EM00316A.

(16) Brown, K. I.; Graham, K. E.; Boehm, A. B. Risk-Based Threshold of Gull-Associated Fecal Marker Concentrations for Recreational Water. *Environ. Sci. Technol. Lett.* **2017**, *4* (2), 44–48. https://doi.org/10.1021/acs.estlett.6b00473.

(17) USEPA. *Recreational Water Quality Criteria*; OFFICE OF WATER 820-F-12-058; 2012.

(18) Boehm, A. B.; Silverman, A. I.; Schriewer, A.; Goodwin, K. Systematic Review and Meta-Analysis of Decay Rates of Waterborne Mammalian Viruses and Coliphages in Surface Waters. *Water Res.* **2019**, 114898. https://doi.org/10.1016/j.watres.2019.114898.

(19) Nappier, S. P.; Hong, T.; Ichida, A.; Goldstone, A.; Eftim, S. E. Occurrence of Coliphage in Raw Wastewater and in Ambient Water: A Meta-Analysis. *Water Res.* **2019**, *153*, 263–273. https://doi.org/10.1016/j.watres.2018.12.058.

(20) Soller, J. A.; Eisenberg, J. N. S. An Evaluation of Parsimony for Microbial Risk Assessment Models. *Environmetrics* **2008**, *19* (1), 61–78.

(21) Yau, V.; Wade, T. J.; de Wilde, C. K.; Colford Jr., J. M. Skin-Related Symptoms Following Exposure to Recreational Water: A Systematic Review and Meta-Analysis. *Water Qual Expo Health* **2009**, *1*, 79–103.

(22) Soller, J. A.; Schoen, M.; Steele, J. A.; Griffith, J. F.; Schiff, K. C. Incidence of Gastrointestinal Illness Following Wet Weather Recreational Exposures: Harmonization of Quantitative Microbial Risk Assessment with an Epidemiologic Investigation of Surfers. *Water Res.* **2017**, *121*, 280–289. https://doi.org/10.1016/j.watres.2017.05.017.

(23) Soller, J. A.; Eftim, S.; Wade, T. J.; Ichida, A. M.; Clancy, J. L.; Johnson, T. B.; Schwab, K.; Ramirez-Toro, G.; Nappier, S.; Ravenscroft, J. E. Use of Quantitative Microbial Risk Assessment to Improve Interpretation of a Recreational Water Epidemiological Study. *Microb. Risk Anal.* **2015**. https://doi.org/10.1016/j.mran.2015.04.001.

(24) Schoen, M. E.; Ashbolt, N. J. Assessing Pathogen Risk to Swimmers at Non-Sewage Impacted Recreational Beaches. *Environ. Sci. Technol.* **2010**, *44*, 2286–2291.

(25) Soller, J. A.; Eftim, S. E.; Warren, I.; Nappier, S. P. Evaluation of Microbiological Risks Associated with Direct Potable Reuse. *Quant. Microb. Risk Assess. Reclaimed Water* **2017**, *5*, 3–14. https://doi.org/10.1016/j.mran.2016.08.003.

(26) DeFlorio-Barker, S.; Arnold, B. F.; Sams, E. A.; Dufour, A. P.; Colford Jr, J. M.; Weisberg, S. B.; Schiff, K. C.; Wade, T. J. Child Environmental Exposures to Water and Sand at the Beach: Findings from Studies of over 68,000 Subjects at 12 Beaches. *J. Expo. Sci. Environ. Epidemiol.* **2018**, *28*, 93–100.

(27) Colford, J. M., Jr.; Wade, T. J.; Schiff, K. C.; Wright, C. C.; Griffith, J. F.; Sandhu, S. K.; Burns, S.; Sobsey, M.; Lovelace, G.; Weisberg, S. B. Water Quality Indicators and the Risk of Illness at Beaches with Nonpoint Sources of Fecal Contamination. *Epidemiology* **2007**, *18* (1), 27–35.

(28) Shanks, O. C.; Kelty, C. A.; Sivaganesan, M.; Varma, M.; Haugland, R. A. Quantitative PCR for Genetic Markers of Human Fecal Pollution. *Appl. Environ. Microbiol.* **2009**, *75* (17), 5507–5513.
(29) Shanks, O. C.; White, K.; Kelty, C. A.; Sivaganesan, M.; Blannon, J.; Meckes, M.; Varma, M.; Haugland, R. A. Performance of PCR-Based Assays Targeting Bacteroidales Genetic Markers of Human Fecal Pollution in Sewage and Fecal Samples. *Environ. Sci. Technol.* **2010**, *44* (16), 6281–6288.
(30) Maier, R. M.; Pepper, I. L.; Gerba, C. P. *Environmental Microbiology*, 2nd ed.; Elsevier: New York, 2009.
(31) Messner, M. J.; Berger, P.; Nappier, S. P. Fractional Poisson--a Simple Dose-Response Model for Human Norovirus. *Risk Anal.* **2014**, *34* (10), 1820–1829.
(32) Medema, G. J.; Teunis, P. F. M.; Havelaar, A. H.; Haas, C. N. Assessment of the Dose-Response Relationship of Campylobacter Jejuni. *Risk Anal. Prod. Safe Food* **1996**, *30* (1), 101–111. https://doi.org/10.1016/0168-1605(96)00994-4.
(33) Rose, J. B.; Gerba, C. P. Use of Risk Assessment for Development of Microbial Standards. *Water Sci. Technol.* **1991**, *24*, 29–34.
(34) Messner, M. J.; Berger, P. Cryptosporidium Infection Risk: Results of New Dose-Response Modeling. *Risk Anal.* **2016**, *36* (10), 1969–1982. https://doi.org/10.1111/risa.12541.
(35) Soller, J. A.; Bartrand, T.; Ashbolt, N. J.; Ravenscroft, J.; Wade, T. J. Estimating the Primary Aetiologic Agents in Recreational Freshwaters Impacted by Human Sources of Faecal Contamination. *Water Res.* **2010**, *44* (16), 4736–4747.
(36) Wade, T. J.; Augustine, S. A. J.; Griffin, S. M.; Sams, E. A.; Oshima, K. H.; Egorov, A. I.; Simmons, K. J.; Eason, T. N.; Dufour, A. P. Asymptomatic Norovirus Infection Associated with Swimming at a Tropical Beach: A Prospective Cohort Study. *PLOS ONE* **2018**, *13* (3), e0195056. https://doi.org/10.1371/journal.pone.0195056.
(37) Koivunen, J.; Siitonen, A.; Heinonen-Tanski, H. Elimination of Enteric Bacteria in Biological–Chemical Wastewater Treatment and Tertiary Filtration Units. *Water Res.* **2003**, *37* (3), 690–698. https://doi.org/10.1016/S0043-1354(02)00305-6.
(38) Lemarchand, K.; Lebaron, P. Occurrence of Salmonella Spp. and Cryptosporidium Spp. in a French Coastal Watershed: Relationship with Fecal Indicators. *FEMS Microbiol Letts* **2003**, *218* (1), 203–209.
(39) Haas, C. N.; Rose, J. B.; Gerba, C. P. *Quantitative Microbial Risk Assessment*; Lett. Appl. Microbiol.; J.W. Wiley, Inc., 1999.
(40) Fazil, A. M. A Quantitative Risk Assessment Model for Salmonella, Drexel University, Philadelphia, PA, 1996.
(41) Teunis, P.; Nagelkerke, N. J.; Haas, C. N. Dose-Response Models for Infectious Gastroenteritis. *Risk Anal.* **1999**, *19* (6), 1251–1260.
(42) Stampi, S.; Varoli, O.; Zanetti, F.; De Luca, G. Arcobacter Cryaerophilus and Thermophilic Campylobacters in a Sewage Treatment Plant in Italy: Two Secondary Treatments Compared. *Epidemiol. Infect.* **1993**, *110* (3), 633–639. https://doi.org/10.1017/S0950268800051050.
(43) Teunis, P.; Van den Brandhof, W.; Nauta, M.; Wagenaar, J.; Van den Kerkhof, H.; Van Pelt, W. A Reconsideration of the Campylobacter Dose-Response Relation. *Epidemiol. Infect.* **2005**, *133* (4), 583–592.

(44) García-Aljaro C.; Bonjoch X.; Blanch A.R. Combined Use of an Immunomagnetic Separation Method and Immunoblotting for the Enumeration and Isolation of Escherichia Coli O157 in Wastewaters. *J. Appl. Microbiol.* **2004**, *98* (3), 589–597. https://doi.org/10.1111/j.1365-2672.2004.02497.x.

(45) Teunis, P.; Ogden, I. D.; Strachan, N. J. C. Hierarchical Dose Response of E. Coli O157:H7 from Human Outbreaks Incorporating Heterogeneity in Exposure. *Epidemiol. Infect.* **2008**, *136* (6), 761–770. https://doi.org/10.1017/S0950268807008771.

(46) Ludwig, K.; Sarkim, V.; Bitzan, M.; Karmali, M. A.; Bobrowski, C.; Ruder, H.; Laufs, R.; Sobottka, I.; Petric, M.; Karch, H.; et al. Shiga Toxin-Producing Escherichia Coli Infection and Antibodies against Stx2 and Stx1 in Household Contacts of Children with Enteropathic Hemolytic-Uremic Syndrome. *J. Clin. Microbiol.* **2002**, *40* (5), 1773–1782. https://doi.org/10.1128/JCM.40.5.1773-1782.2002.

(47) Werber, D.; Mason, B. W.; Evans, M. R.; Salmon, R. L. Preventing Household Transmission of Shiga Toxin–Producing Escherichia Coli O157 Infection: Promptly Separating Siblings Might Be the Key. *Clin. Infect. Dis.* **2008**, *46* (8), 1189–1196. https://doi.org/10.1086/587670.

(48) Bielaszewska, M.; Janda, J.; Bláhová, K.; Minaríková, H.; Jíková, E.; Karmali, M. A.; Laubová, J.; Sikulová, J.; Preston, M. A.; Khakhria, R.; et al. Human Escherichia Coli O157:H7 Infection Associated with the Consumption of Unpasteurized Goat's Milk. *Epidemiol. Infect.* **1997**, *119* (3), 299–305.

(49) Yang Jian; Schneider Orren D.; Jjemba Patrick K.; Lechevallier Mark W. Microbial Risk Modeling for Main Breaks. *J. - Am. Water Works Assoc.* **2015**, *107* (2), E97–E108. https://doi.org/10.5942/jawwa.2015.107.0010.

(50) Crockett, C. S. The Role of Wastewater Treatment in Protecting Water Supplies Against Emerging Pathogens. *Water Environ. Res.* **2007**, *79* (3), 221–232. https://doi.org/10.2175/106143006X111952.

(51) Harwood, V. J.; Levine, A. D.; Scott, T. M.; Chivukula, V.; Lukasik, J.; Farrah, S. R.; Rose, J. B. Validity of the Indicator Organism Paradigm for Pathogen Reduction in Reclaimed Water and Public Health Protection. *Appl. Environ. Microbiol.* **2005**, *71* (6), 3163–3170. https://doi.org/10.1128/aem.71.6.3163-3170.2005.

(52) Nasser, A. M. Removal of Cryptosporidium</Em> by Wastewater Treatment Processes: A Review. *J. Water Health* **2016**, *14* (1), 1. https://doi.org/10.2166/wh.2015.131.

(53) Schoen, M. E.; Ashbolt, N. J.; Jahne, M. A.; Garland, J. Risk-Based Enteric Pathogen Reduction Targets for Non-Potable and Direct Potable Use of Roof Runoff, Stormwater, and Greywater. *Quant. Microb. Risk Assess. Reclaimed Water* **2017**, *5*, 32–43. https://doi.org/10.1016/j.mran.2017.01.002.

(54) US EPA. *National Primary Drinking Water Regulations: Long Term 2 Enhanced Surface Water Treatment Rule (LT2ESWTR)*; Final Rule. 40CFR Parts 9, 141 and 142, volume 71, Number 654; 2006.

(55) Kitajima, M.; Haramoto, E.; Iker, B. C.; Gerba, C. P. Occurrence of Cryptosporidium, Giardia, and Cyclospora in Influent and Effluent Water at Wastewater Treatment Plants in Arizona. *Sci. Total Environ.* **2014**, *484*, 129–136. https://doi.org/10.1016/j.scitotenv.2014.03.036.

(56) Eisenberg, J. N. S.; Seto, E. Y. W.; Olivieri, A. W.; Spear, R. C. Quantifying Water Pathogen Risk in an Epidemiological Framework. *Risk Anal* **1996**, *16* (4), 549–563.
(57) Eftim, S. E.; Hong, T.; Soller, J.; Boehm, A.; Warren, I.; Ichida, A.; Nappier, S. P. Occurrence of Norovirus in Raw Sewage – A Systematic Literature Review and Meta-Analysis. *Water Res.* **2017**, *111* (Supplement C), 366–374. https://doi.org/10.1016/j.watres.2017.01.017.
(58) Teunis, P. F.; Moe, C. L.; Liu, P.; S, E. M.; Lindesmith, L.; Baric, R. S.; Le Pendu, J.; Calderon, R. L. Norwalk Virus: How Infectious Is It? *J Med Virol* **2008**, *80* (8), 1468–1476.
(59) Hurst, C. J.; McClellan, K. A.; Benton, W. H. Comparison of Cytopathogenicity, Immunofluorescence and In Situ DNA Hybridization as Methods for the Detection of Adenoviruses. *Water Res.* **1988**, *22* (12), 1547–1552. https://doi.org/10.1016/0043-1354(88)90167-4.
(60) Hewitt, J.; Leonard, M.; Greening, G. E.; Lewis, G. D. Influence of Wastewater Treatment Process and the Population Size on Human Virus Profiles in Wastewater. *Water Res.* **2011**, *45* (18), 6267–6276. https://doi.org/10.1016/j.watres.2011.09.029.
(61) TEUNIS, P.; SCHIJVEN, J.; RUTJES, S. A Generalized Dose-Response Relationship for Adenovirus Infection and Illness by Exposure Pathway. *Epidemiol. Infect.* **2016**, *144* (16), 3461–3473. https://doi.org/10.1017/S0950268816001862.

Table 1. Untreated sewage concentrations for reference enteric pathogens and coliphage, and dose-response relations, and $P_{ill|inf}$ for reference enteric pathogens. Unit (refs) is the concentration in sewage, μ is the dose, P_{inf} is probability of infection, $P_{ill|inf}$ is probability of becoming ill after infection. Note that units of pathogens is per liter and for coliphage is per ml to reflect the units used in the literature for these parameters. CFU is colony forming unit, MPN is most probable number, copy refers to gene copy number, IU is infectious unit, and PFU is plaque forming unit. ${}_1F_1$ is the hypergeometric function. When specified, $P_{ill|inf}$ are represented by a range of parameters, as indicated, drawn from a uniform distribution. $P_{ill|inf}$ for *Campylobacter* is dose-dependent with $r = 2.44 \times 10^8$ and $\nu = 3.63 \times 10^{-9}$. References (Refs) for P_{inf} and $P_{ill|inf}$ are provided in the last column. References for sewage concentration range are provided adjacent to the unit. # The two values separated by a comma are the minimum and maximum of the $\log_{10}$-uniform distribution. ^ The two values separated by a comma are the mean and standard deviation of a $\log_{10}$-normal distribution. * Lower range is not detected and -1 is used as a lower bound. NA means not applicable. The coliphage distributions reflect those numbers reported by Nappier et al.[19] from North America, accounting for geographic variability.

Organism/Target	C_{i_sewage}	Unit (refs)	P_{inf}	$P_{ill\|inf}$ (distribution)	Refs
Salmonella spp.	[0.5,5]#	CFU/L [37,38]	$1-(1+\mu/2884)^{-0.3126}$	0.17-0.4 (uniform)	[39–41]
Campylobacter	[2.9,4.6] #	MPN/L[42]	$1-\ {}_1F_1(0.024,0.024+0.011,-\mu)$	$1-(1+\nu\mu)^{-r}$	[43]
E. coli O157:H7	[-1,3.3] #,*	CFU/L [44]	$1-(1+\mu/48.8)^{-0.248}$	0.2-0.6 (uniform)	[45–48]
Cryptosporidium	[-0.52, 3.7] #	oocysts/L[49–53]	$1 - \exp(-0.09\ \mu)$	0.3-0.7(uniform)	[54]
Giardia	[0.51,4.2] #	cysts/L [51,55]	$1 - \exp(-0.0199\ \mu)$	0.2-0.7 (uniform)	[33,56]
norovirus	[4.0,1.1] ^	copy/L[57]	$1-\ {}_1F_1\ (0.04,\ 0.04+\ 0.055,\ -\mu)$	0.3-0.8 (uniform)	[58]
adenovirus	[1.75,3.84] #	IU/L[25,59,60]	$1-\ {}_1F_1\ (5.11,\ 5.11+\ 2.8,\ -\mu)$	0.5 (uniform)	[61]
somatic coliphage	[3.0, 1.3]^	PFU/ml[19]	NA	NA	
F+ coliphage	[2.8, 1.0]^	PFU/ml[19]	NA	NA	

	Temperature = 15°C		Temperature = 25°C	
Organism/ Indicator	**mean $\log_{10}$-*k***	**SD $\log_{10}$-*k***	**mean $\log_{10}$-*k***	**SD $\log_{10}$-*k***
norovirus	-1.22	0.13	-0.80	0.18
adenovirus	-1.45	0.14	-0.57	0.23
Salmonella	-0.22	0.059	-0.039	0.082
Campylobacter	0.27	0.12	0.63	0.25
E. coli O157:H7	-0.43	0.038	-0.28	0.061
Giardia	-1.14	0.21	-2.04	0.29
Cryptosporidium	-1.45	0.14	-0.57	0.23
somatic coliphage	-0.26	0.11	0.34	0.18
F+ coliphage	-0.24	0.06	0.31	0.084

Table 2. Normal distributions of $\log_{10}$-transformed *k* for the organisms and indicators used in the QMRA. Mean and standard deviation (SD) are provided. Units of *k* before they were $\log_{10}$-transformed are d^{-1}.

Indicator	T (°C)	b_1	m_1	b_2	m_2
Somatic coliphage	15	3.25	-0.27	5	-0.239
Somatic coliphage	25	3.30	-1.30	5	-0.950
F+ coliphage	15	3.35	-0.254	4.8	-0.250
F+ coliphage	25	3.42	-0.98	4.8	-0.887

Table 3. Slope and intercept for the following linear relationship: $\log_{10}C_{meas|max} = m*\tau_{max} + b$. This line describes the relationship between the maximum plausible concentration of the indicator in surface waters as a function of contamination age and assumes that the water is 100% raw sewage by volume. The units of $C_{meas|max}$ are PFU/100 mL and the units of τ_{max} are days. b_1 and m_1 were derived empirically by determining when more than 10% of the random Monte Carlo draws for a specified τ and C_{meas} required that the volume fraction of sewage be greater than 1. b_2 and m_2 represent median values as described in the methods section.

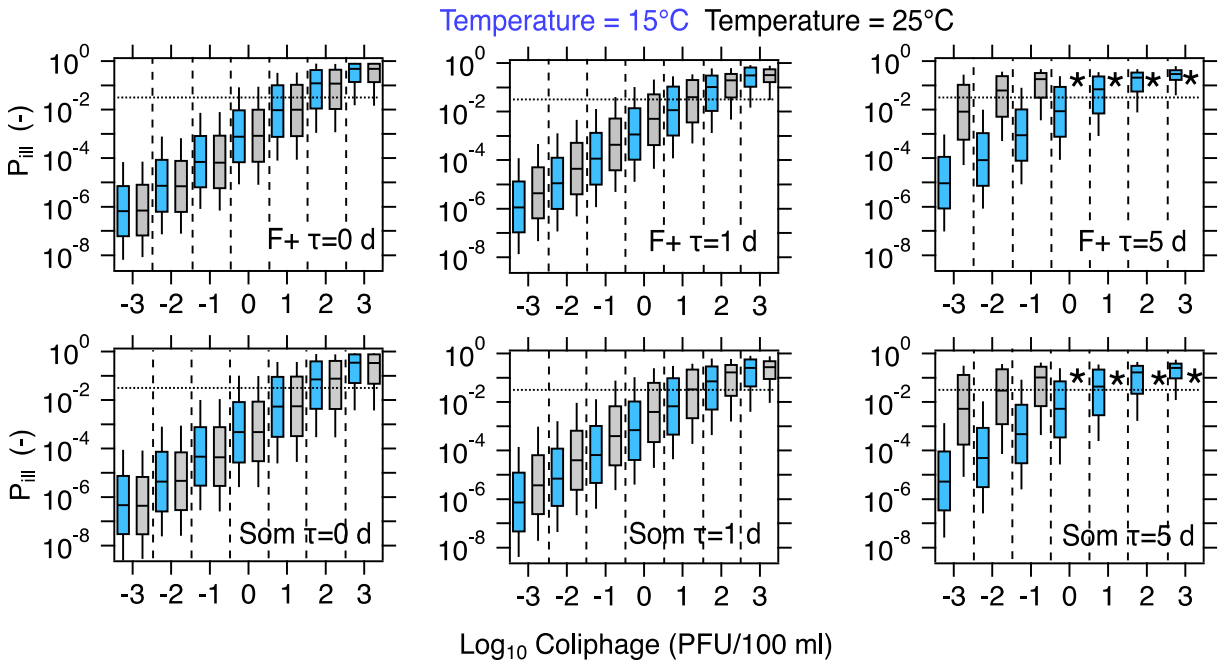


Figure 1. Simulated probability of illness for a subset of the modeled concentrations of somatic (Som) and F+ coliphage in surface waters (C_{meas}). The box and whiskers illustrate the median of the 10,000 simulations (horizontal line through the middle of the box), the 25th and 75th percentiles of the simulations (bottom and top of the box, respectively) and the 10th and 90th percentiles (top and bottom of the whiskers). The results for 3 different ages are shown for both temperatures. Results are not shown for C_{meas} if at the specified τ, 10,000 iterations with F<1 could not be obtained from 15,000 random draws (see methods); a * is shown where that occurs. The horizontal line shows P_{ill} = 32 /1000.

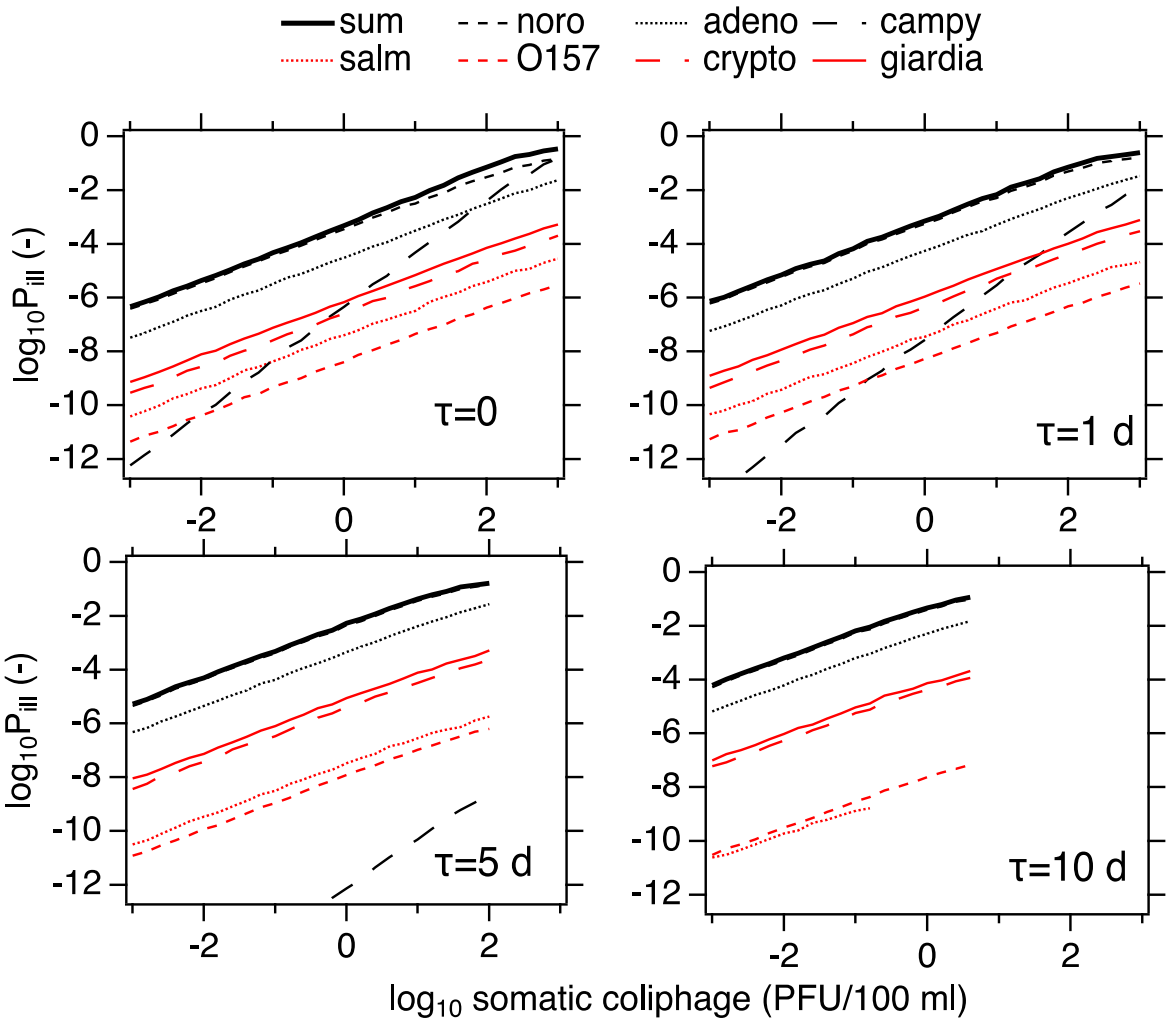


Figure 2. The median probability of illness from exposure to each considered pathogen given C_{meas} for somatic coliphage at T=15°C for a subset of ages considered. "sum" refers to the cumulative risk. noro, adeno, campy, salm, O157, crypto, and giardia refer to the probability of illness attributable to exposure to norovirus, adenovirus, *Campylobacter*, *Salmonella*, *E. coli* O157:H7, *Cryptosporidium*, and *Giardia,* respectively. Results are not shown for C_{meas} that are unrealistic given the age of contamination τ (where F≥1 for too many of the Monte Carlo draws, as explained in the caption of Figure 1 and in the text). Results for Campylobacter are not shown for τ=10 d because the probability of illness is essentially 0 and cannot be easily displayed on a plot with a log-transformed y-axis.

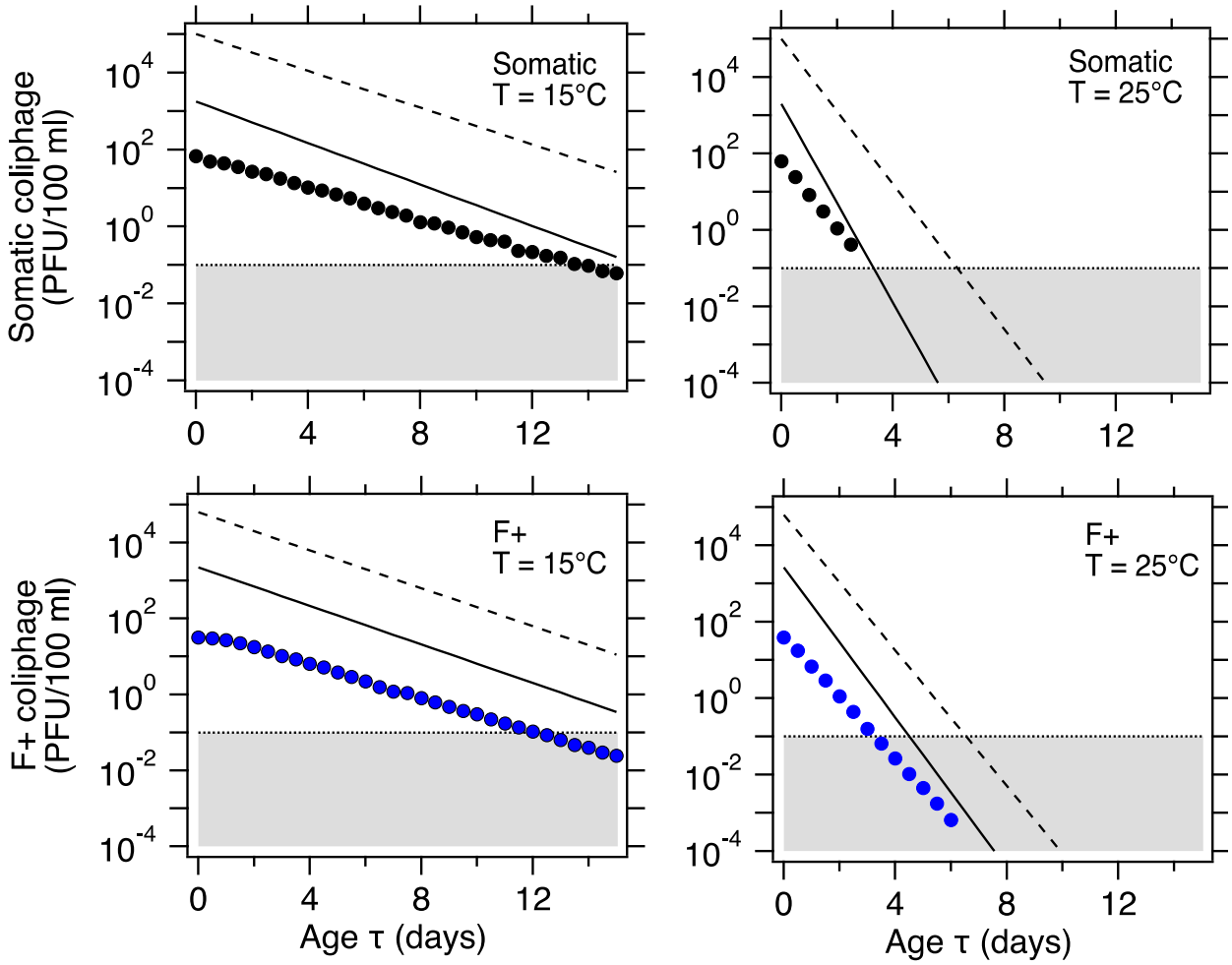


Figure 3. Risk-based thresholds of coliphage as a function of the age τ of contamination for waters at 15°C and 25°C (filled circle markers). The risk-based threshold is the concentration of indicator at which the median simulated risk is 32/1000. The lines represent $C_{meas|max}$ as a function of τ when b_1 and m_1 are used (C_{max1}, solid lines) and when b_2 and m_2 are used (C_{max2}, dashed lines) as presented in Table 3. The risk-based threshold is not calculated when it exceeds the C_{max1}, but it can be readily estimated by extending the linear line to greater τ. Somatic is somatic coliphage (black), F+ is male specific F+ coliphage (blue). The grey shaded area of the plots shows concentrations lower than 0.1 PFU/100 ml (or 1 PFU/liter) which is generally a lowest detectable concentration of standard coliphage enumeration assays.

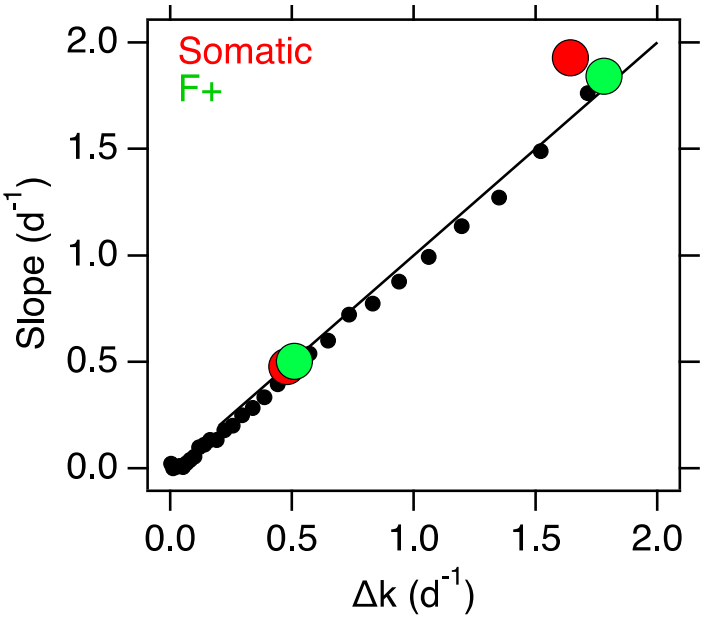


Figure 4. The slope of the line ("slope") relating the change in the natural-log (ln) transformed risk-based threshold as a function of τ versus Δk (median of $k_{indicator}$-k_{noro}). The black markers illustrate simulated values obtained by systematically changing $k_{indicator}$ across a range of values. The colored markers show results from the QMRA simulations presented in Figure 3). The colored circles with larger Δk are the results at 25°C while the ones with smaller Δk are results at 15°C. The solid line shows the 1:1 line.

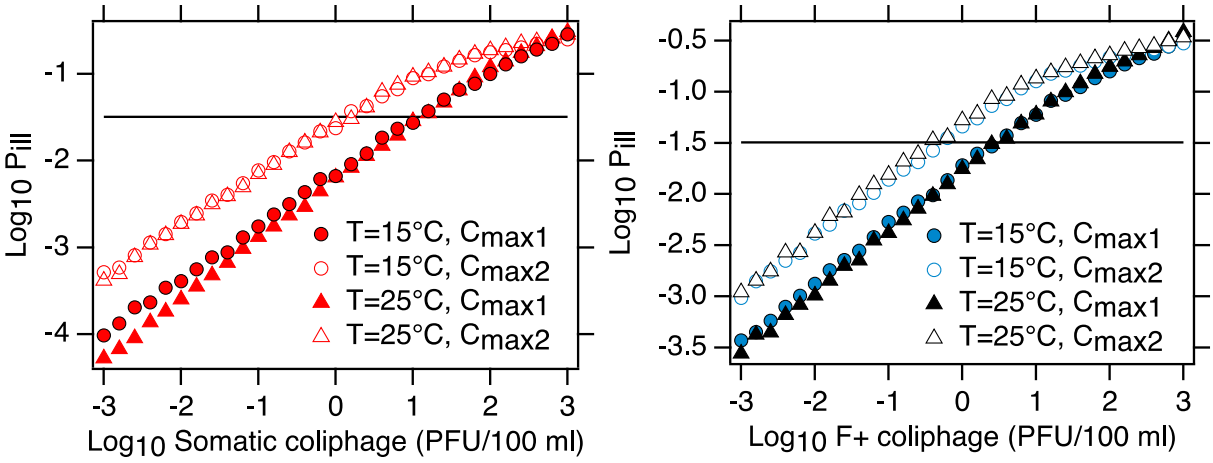
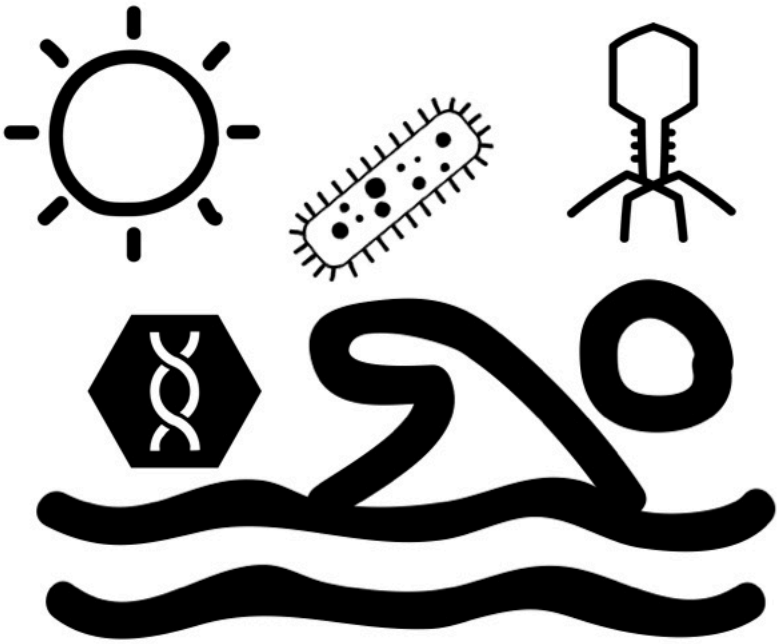


Figure 5. Median simulated $\log_{10}$-transformed Pill as a function of coliphage concentration in surface waters when the age of contamination is unknown and τ is allowed to vary according to a uniform distribution defined by 0 and τ_{max}. τ_{max} is determined from one of the two relationships presented in Table 3 where C_{max1} uses b_1 and m_1, and C_{max2} uses b_2 and m_2



A quantitative microbial risk assessment framework is used to derive risk-based surface water quality thresholds for coliphages.