A somatic coliphage threshold approach to improve the management of activated sludge wastewater treatment plant effluents in resource-limited regions

Luz Chacón[a#], *Kenia Barrantes*[a#], *Carolina Santamaría-Ulloa*[a], *Melissa Solano*[a†], *Liliana Reyes*[a], *Lizeth Taylor*[b¤], *Carmen Valiente*[c], *Erin M. Symonds*[d], *Rosario Achí*[a]

Running Tittle: Somatic coliphage threshold for WWTP effluent management

[a]Health Sciences Research Institute (Instituto de Investigaciones en Salud (INISA)), Universidad de Costa Rica, Montes de Oca, Costa Rica

[b]College of Microbiology (Facultad de Microbiología), Universidad de Costa Rica, Montes de Oca, Costa Rica

[c]National Water Laboratory, Costa Rican Institute of Aqueducts and Sewerage (Laboratorio Nacional de Aguas, Instituto Costarricense de Acueductos y Alcantarillados), Tres Ríos, Costa Rica

[d]College of Marine Science, University of South Florida, 140 Seventh Avenue South, St. Petersburg, Florida 33701, United States of America

[¤]Deceased 13 December 2013

[#]Luz Chacón luz.chacon@ucr.ac.cr /[#]Kenia Barrantes, kenia.barrantes@ucr.ac.cr

Present Addresses

[†]College of Microbiology (Facultad de Microbiología), Universidad de Costa Rica, Montes de Oca, Costa Rica

Keywords: wastewater treatment, enteric viruses, somatic coliphage, fecal indicator bacteria, activated sludge, threshold values

Abstract

Effective wastewater management is crucial to ensure the safety of water reuse projects and effluent discharge into surface waters. Multiple studies have demonstrated that municipal wastewater treatment with conventional activated sludge processes is inefficient for the removal of the wide spectrum of viruses in sewage. In this study, a well-accepted statistical approach was used to investigate the relationship between viral indicators and human enteric viruses during wastewater treatment in a resource-limited region. Influent and effluent samples from five urban wastewater treatment plants (WWTP) in Costa Rica were analyzed for somatic coliphage and human enterovirus, hepatitis A virus, norovirus genotype I and II, and rotavirus. All WWTP provide primary treatment followed by conventional activated sludge treatment prior to discharge into surface waters that are indirectly used for agricultural irrigation. The results revealed a statistically significant relationship between the detection of at least one of the five human enteric viruses and somatic coliphage. Multiple logistic regression and Receiver Operating Characteristic curve analysis identified a threshold of 3.0×10^3 (3.5-log_{10}) somatic coliphage plaque forming unit per 100 mL, which corresponded to an increased likelihood of encountering enteric viruses above the limit of detection ($>1.83\times10^2$ virus target/100 mL). Additionally, quantitative microbial risk assessment was executed for famers indirectly reusing WWTP effluent that met the proposed threshold. The resulting estimated median cumulative annual disease burden complied with World Health Organization recommendations. Future studies are needed to validate the proposed threshold for use in Costa Rica and other regions.

47 **Importance**

48 Effective wastewater management is crucial to ensure safe direct and indirect water
49 reuse; nevertheless, few countries have adopted the virus log reduction value management
50 approach established by the World Health Organization. In this study, we investigated an
51 alternative and/or complementary approach to the virus log reduction value framework for the
52 indirect reuse of activated sludge treated wastewater effluent. Specifically, we employed a
53 well-accepted statistical approach to identify a statistically sound somatic coliphage threshold
54 value, which corresponded to an increased likelihood of human enteric virus detection. This
55 study demonstrates an alternative approach to the virus log reduction value framework, which
56 can be applied to improve wastewater reuse practices and effluent management.
57

1. Introduction

Conventional activated sludge is an aerobic, secondary wastewater treatment technology that takes advantage of biological processes to remove organic matter and is commonly used in low-, middle- and high-income countries (1). Frequently, activated sludge wastewater treatment plant (WWTP) effluent does not receive additional treatment, even though it is well-known that pathogen removal can be insufficient for safe water reuse (2–9). This is particularly true for enteric viruses because traditional activated sludge treatment typically removes viruses 2.02-$\log_{10}$ (1, 10). Currently, human enteric viruses cause a significant fraction of the disease burden related to wastewater pollution worldwide. Direct and indirect wastewater reuse (e.g., agricultural irrigation, recreational activities in contaminated surface waters) represents a public health risk; thus, the microbial quality of WWTP effluent should be monitored to manage those risks (11, 12).

Fecal indicator bacteria (e.g., fecal coliform, enterococci, and *Escherichia coli*) are the most commonly used indicators for assessing WWTP effluent microbial quality (13). They were initially introduced as indicators when *Salmonella* Typhi was the principal pathogen of concern. Despite their effectiveness for indicating bacterial pathogens, several studies have demonstrated that fecal indicator bacteria did not correlate with enteric viruses in WWTP effluent (14–17). Furthermore, high enteric virus concentrations were detected when fecal indicator bacteria concentrations were low.

While fecal indicator bacteria are not useful viral indicators of wastewater treatment processes (18, 19), country-specific legislation concerning WWTP effluent reuse and discharge frequently rely on fecal indicator bacteria (13). No universally accepted viral indicator or criteria exists to date (10). Some governments now include viral indicators, either human reference viral

pathogens, somatic coliphage or F+ coliphage, to determine WWTP virus reductions (summarized in (7)). Meta-analyses conducted in wastewater matrices report bacteriophages, particularly somatic coliphage, as good surrogates of human enteric viruses because of their similar characteristics, high concentrations, and low-cost methods that distinguish infectious viruses (10, 20, 21)

Currently, the World Health Organization (WHO) recommends a multiple-barrier approach to managing WWTP effluent, in which a reference human enteric virus log reduction value is associated with each treatment process (13). Practitioners define the physical and chemical conditions that achieve the target virus log reduction value, and then assume that the log reduction value remains constant if the physical-chemical conditions do not change (22). While this approach was accepted among experts, most countries in the world have yet to apply this management approach for a variety of reasons (7). Even though routine monitoring is not required if physical-chemical conditions remain constant, this log reduction value effluent management approach has been met with resistance in many countries because it is difficult to implement into practice given that it is not a threshold value.

Additionally, the reference human enteric virus analyses required to identify the conditions associated with a target log reduction value are not feasible for many municipal WWTPs in high-income settings, let alone feasible in middle- and low-income contexts. They require expertise and sophisticated laboratory equipment, are time consuming, costly, and enteric virus concentrations are frequently below detectable concentrations (4, 22–24). Furthermore, these reference pathogen analyses are typically executed using molecular methods, which cannot distinguish infectious and non-infectious viruses (7, 23, 25). Even though some countries' legislation focuses on reference enteric virus log reduction values, somatic and F+ coliphage

104 have also been used in the log reduction value management framework (10, 11, 26). Regardless
105 of the human enteric reference virus or indicator used, the log reduction value management
106 framework has been criticized for not effectively protecting public health because it focuses on
107 removal and disregards the variability of human enteric virus concentrations in WWTP influent.
108 Consequently, additional 2-3-$\log_{10}$ removal can be needed to ensure safe WWTP discharge and
109 reuse, even if log reduction value targets are met (4).

110 Prior to the virus log reduction value management approach two decades ago, a somatic
111 coliphage threshold (3-$\log_{10}$ PFU/100 mL) associated with infectious enterovirus concentrations
112 was proposed to better manage WWTP effluent discharges (20). However, this threshold value
113 was never applied to management and needs to be re-calculated because it is based on a non-
114 robust statistical approach and considers just one human enteric virus (27). Given the difficulties
115 and disadvantages associated with applying the virus log reduction value management approach,
116 the objective of this study was to determine a statistically-sound, robust somatic coliphage
117 concentration threshold useful for monitoring WWTP effluents.

118 To demonstrate this approach, somatic coliphage and enteric viruses were monitored at
119 five activated sludge WWTPs in the San José Metropolitan Area, Costa Rica. The human enteric
120 virus included in this study were human enterovirus (EV), hepatitis A virus (HAV), norovirus
121 genotype I and II (NoVGI and NoVGII), and rotavirus group A (RV) because they are an
122 important cause of outbreaks and diarrheal illness in Costa Rica (28, 29). Data were analyzed
123 using the most-accepted, robust statistical methods (multiple logistic regression models and
124 receiver operating characteristics (ROC) curves (27, 30, 31) to establish a useful threshold that
125 corresponds to the minimum somatic coliphage concentration associated with increased human
126 enteric virus detection. Since Costa Rican domestic WWTP effluent is currently managed using

fecal coliform concentration thresholds that vary based upon potential wastewater reuse activities, fecal coliforms were also monitored simultaneously and similar statistical analyses were executed to compare the current bacterial indicator with the proposed viral indicator. Finally, quantitative microbial risk assessment was used to estimate the annual disease burden associated with indirectly irrigating with WWTP effluent that met the proposed somatic coliphage threshold.

2. Materials and Methods

2.1 Wastewater treatment plant sample collection

A total of 119, 1.5 L influent (n = 60) and effluent (n = 59) samples were collected from five urban WWTPs located in the San José Metropolitan Area, Costa Rica (Figure 1) All of the WWTPs are small in size (i.e., treating waste from 123 to 1033 inhabitants and only receive domestic wastewater) (5, 6, 32). They consist of primary treatment followed by secondary treatment via conventional activated sludge processes. The WWTP effluents are discharged into the Virilla River, which are also source water for agricultural irrigation. None of these wastewater treatment facilities disinfect effluent prior to surface water discharge. Since this study was executed in a tropical country, there are two seasons: (1) the dry season from December through April and (2) the rainy season from May through November. In order to account for seasonal differences in weather and human enteric virus seasonality, grab samples were collected from each WWTP between 9:00 a.m. and 12:00 p.m. on three consecutive days, for each of the following months in 2013: March, May, October, and December. All samples were collected in sterile, amber bottles and maintained at 4 °C until processed. All samples were

150 analyzed for somatic coliphages and fecal coliform concentrations. Presence/absence analyses
151 for the following human enteric viruses were carried out on a subset of samples using PCR-based
152 methods: EV (n = 117), HAV (n = 117), norovirus GI (NoVGI; n = 72) and GII (NoVGII; n =
153 72); and RV (n = 79).
154
155 ***2.2 Fecal coliform analyses***
156 Fecal coliforms most probable number (MPN) concentrations were determined by
157 multiple- tube fermentation (MPN/100 mL) according to Method 9221E within 8 h of collection
158 (33). Briefly, all samples were inoculated in a series of five tubes with lauryl sulfate broth, in
159 which the WWTP influent and effluent samples were serially diluted to a concentration of
160 1:1,000,000 and 1:100,000, respectively, prior to inoculation. Confirmation was executed after
161 48 ± 4 h of incubation at 35 °C, an inoculum of each tube with bacterial growth and gas were
162 transferred to EC-MUG broth and were incubated for 24 ± 2 h at 44.5 °C; tubes positive for fecal
163 coliforms had bacterial growth and gas characteristics. A positive control (*E. coli* ATCC 25922),
164 a negative control (*Salmonella spp*. ATCC 13076), and a blank (containing the dilution buffer as
165 inoculate) were analyzed alongside all samples. No contamination was observed, and all positive
166 and negative controls generated positive and negative results, respectively.
167
168 ***2.3 Wastewater Pre-treament for virus isolation and concentration***
169 All samples were pre-filtered with a metal sieve (0.15 mm pore) in order to break up
170 large organic particles. Viruses were concentrated in accordance with, and following, the
171 Standard Methods for the Examination Water and Wastewater (Section 9510C; (33)). Briefly, the
172 pre-filtered wastewater sample (1.25 L) was successively filtered through three filters pretreated

with 3% beef extract (pH 7.2; Oxoid®, United Kingdom) to remove larger particles and prevent viruses from sticking to the filters: (1) 47 mm, 80 µm glass fiber filter (13400-47-Q; Sartorius®, Germany); (2) 47 mm, 1.2 µm nitrate cellulose filter (11303-47-N; Sartorius®, Germany); and (3) 47 mm, 0.4 µm acetate cellulose filter (11106-47-ACN; Sartorius®, Germany). This filtrate was divided into two parts: 250 mL for somatic coliphage analyses and 1 L for enteric virus analyses. With the exception of the somatic coliphage analyses for WWTP effluent, the filtrate was stored at -70 ºC prior to human enteric virus concentration and WWTP influent somatic coliphage quantification.

2.4 Human enteric virus concentration and detection

One liter of filtered WWTP influent and effluent was concentrated using a modified adsorption-elution method (Method 9510B) (33). Sample pH was adjusted to 3.5 with HCl (0.1 N) and filtered with 47 -mm, 0.2 -nm cellulose acetate filter (1110tr-47N Sartorius®, Germany) to adsorb the viruses onto the filter; approximately three filters were used for each sample in order to filter the entire 1 -L sample. Subsequently, the viruses were eluted off the filter(s) with 15 mL beef extract 3% pH 9.0. All eluate was collected and precipitated at 4 ºC with PEG8000 and 17.5 g/L NaCl (34). The final virus concentrates (0.5 ml) were stored at -70 °C prior to RNA purification. The concentration efficiency of this method ranged between 40% - 90% in previous studies (33, 35). It was also tested with a poliovirus vaccine strain (Sabin vaccine strain), in which the concentration of the original and concentrated samples were determined using the Dulbecco plates method (36) with Hep-2 cells. The concentrated sample was 1-$\log_{10}$ more concentrated in comparison to the original sample (data not shown).

195 Viral RNA (50 µl) was obtained from the entire final virus concentrate (0.5 ml) using the
196 NucleoSpin RNA Virus kit (Macherey Nagel®, Germany) and cDNA (20 µl) was synthesized
197 from 8.0 µl viral RNA using the RevertAid™ H Minus First Strand cDNA Synthesis kit with
198 random hexamers (Thermo Scientific®, USA), both following the manufacturer's instructions.
199 Presence/absence analyses for the following human enteric viruses were carried out on a subset
200 of samples using reverse transcriptase polymerase chain reaction (RT-PCR)-based methods and
201 previously published assays and conditions (Table 1; (37–40): enterovirus (EV; n = 117),
202 hepatitis A virus (HAV; n = 117), norovirus GI (NoVGI; n = 72) and GII (NoVGII; n = 72), and
203 rotavirus group A (RV; n = 79). Presence-absence human enteric virus data were generated in
204 this study because previous studies demonstrated a better correlation between enteric virus
205 presence/absence and coliphages in comparison with correlations with quantitative enteric virus
206 data (27, 41). All RT-PCR-based analyses were executed using Master Mix 2X (Fermentas®,
207 USA) with a final reaction volume of 25 µL.
208 For the end-point RT-PCR assays (EV and HAV), the Applied BioSystem® Veriti 9902
209 thermocycler was used. A sample was identified as positive when PCR products with the
210 anticipated size (EV, 113 bp; HAV, 266 bp) were visualized using 2% agarose gel
211 electrophoresis with GelRed®. For NoVGI, NoVGII, and RV presence/absence was determined
212 using RT-quantitative PCR (RT-qPCR) with a StepOne Real-Time PCR thermocycler (Applied
213 Biosystems®). A sample was identified as positive if the Cq value was less than 35. For samples
214 with a Cq value greater than 35 and less than 40, the sample was re-run and all samples with
215 mean Cq values ≤ 35 were classified as positive.
216 In addition to a negative control (sterile water), the following positive controls, specific to
217 each assay, were used for each instrument run: RV-, NoVGI-, and NoVGII-positive fecal

samples (Costa Rican National Children's Hospital), the Sabin 1 (NIBSC 1/528) vaccine strain for EV (University of Costa Rica, Department of Microbiology, Virology Section), and HAX-70 strain for HAV (University of Costa Rica, Department of Microbiology, Virology Section). All positive controls yielded positive results and all negative controls were negative. The enteric virus theoretical process detection limit (copies/100 mL) was back-calculated using the following equation, which took into account the efficiency published for each step in the molecular analyses as well as the concentration methods used (Eq. 1):

$$limit\ of\ detection\ \frac{copies}{ml} = \frac{c}{v_1} \times \frac{V_1}{E_1} \times \frac{V_2}{v_2 \times E_2} \times \frac{V_3}{v_3} \times \frac{V_4}{v_4} \times \frac{1}{V_5 \times E_3}$$

where c equals copies that could be detected per RT-qPCR reaction (*i.e.,* lowest copy number detected divided by 2 (difference between double-stranded standard curve material and single-stranded viral RNA)); v_1 equals the volume of cDNA added to the qPCR reaction (5 µl); V_1 equals the total volume of cDNA synthesized (20 µl); E_1 equals the worst-case RT efficiency previously reported (19%; (42)); v_2 equals the volume of RNA in the RT reaction (8 µl); V_2 equals the total volume of RNA purified (50 µl); E_2 equals the worst-case viral RNA purification efficiency (90%; (43)); v_3 equals the volume of PEG concentrate that RNA was purified from (500 µl); V_3 equals the total volume of PEG concentrate (500 µl); v_4 equals the eluate volume that was PEG concentrated (45 mL); V_4 equals the total volume of eluate (45 mL); V_5 equals the total volume of wastewater (1000 mL); and E_3 equals the estimated virus concentration efficiency (40%; (35)). The limit of detection for the assays could have been as few as 10 copies (J. Nordgren, personal communication) and great as 1,000 copies (37, 38). Since the limit of detection of each assay was not tested in this study, the limit of detection (c) was defined as 10 copies and 1,000 copies. Thus, the theoretical process limit of enteric virus detection for any given assay was estimated to range from 183 virus copies/100 mL to 18,300 copies/100 mL.

240

241 ***2.5 Somatic coliphage quantification***

242 Somatic coliphage concentrations were determined according to Methods 9924B Somatic
243 Coliphage Assay and 9924E Single-Agar-Layer Method with modifications: 250 mL sample
244 volumes were filtered with 0.2 µm filter (cellulose acetate, 11107- 91 47N Sartorius®, Germany)
245 that was pretreated with 3% beef extract pH 7.2 (33, 44). Somatic coliphage concentrations were
246 identified in WWTP effluent samples using single-layer plaque assay (undiluted sample) and in
247 WWTP influent samples using double-layer plaque assay (1:10,000 serial-dilution of sample).
248 Analyses used the host strain *E. coli* ATCC 13706. Positive (PhiX174 ATCC 13706-B1 phage)
249 and negative (buffer only) controls were run alongside samples. No contamination was observed,
250 and all controls gave anticipated results.

251

252 ***2.6 Data analyses: statistics and indicator concentration threshold evaluation***

253 Descriptive statistics (mean and standard deviation) and comparative (two-group
254 comparisons) analyses were executed using R' Version 3.5.3 (www.rproject.org) with the
255 appropriate methods for non-parametric uncensored, as well as right-, and left-censored data
256 from the NADA package (45). Mean and standard deviation were calculated for somatic
257 coliphages for WWTP influent, and excluded WWTP influent concentrations that were 5-$\log_{10}$
258 PFU/100 mL below the average (n = 31). These data were excluded because somatic coliphage
259 were analyzed with culture-based analyses that were likely inhibited by high concentrations of
260 household disinfectants (46).

261 The mean and standard deviation were estimated using the Kaplan Meir method for the
262 following censored data: somatic coliphage WWTP effluent, and fecal coliforms WWTP influent

and effluent. All somatic coliphage WWTP effluent concentrations below the detection limits (<1 PFU/ 100 mL; e.g. left-censored) were conservatively censored to 0.9 PFU/ 100 mL (n = 8). All fecal coliforms concentrations greater than the method detection limits (e.g., right-censored) were censored to one plus the highest detectable concentration (*i.e.*, > 8.2-$\log_{10}$ MPN/ 100 mL for WWTP influent (n = 22) and > 6.2-$\log_{10}$ MPN/ 100 mL for WWTP effluent (n = 11)). The Peto-Prentice test is a non-parametric analysis that is appropriate for censored data. It was used to test the null hypotheses that there was no significant difference in indicator concentrations (somatic coliphage or fecal coliform) between WWTP influents, WWTP effluents, and WWTP influent and effluents combined.

In order to calculate an indicator threshold concentration that corresponds to human enteric virus detection, multiple logistic regression models were created to determine the statistical significance and association between each indicator and any human enteric virus detection for WWTP influent and effluent (41). The positive classification for human enteric virus detection was based upon the detection of any of the five viruses, which reflects the existence of a public health risk if any one of the viruses are detected, and was previously recommended for this type of analysis (27). The multiple logistic regression model equation was defined as (Eq. 2):

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1\chi_1 + \beta_2\chi_2$$

where p was human enteric virus detection (PCR positive/negative; dependent and dichotomic variable), β_0 was the intercept, β_1 and β_2 were the regression parameters, χ_1 was the indicator (either somatic coliphage or fecal coliforms) concentration, and χ_2 was the dichotomic variable for season. Analyses were conducted for WWTP influent and effluent separately. The specific WWTP was a controlled factor in the model. Chi-square and unpaired two-sample t-test analyses

were used to identify significant ($p < 0.05$) differences between the multiple logistic regression model parameters. Since the WWTP influent multiple logistic regression models did not yield statistically significant relationships; subsequent analyses were conducted only on the WWTP effluent models.

For each indicators' WWTP effluent multiple logistic regression model, the area under ROC curves were estimated in order to measure the regression model's ability to discriminate between effluent samples with and without the detection of any human enteric virus pathogens. The ROC curve is a plot of sensitivity (true-positive rate, y-axis) and specificity (false-positive rate, x-axis) of the logistic regression model and it was used predict human enteric virus detection (any of the five human enteric viruses) in effluent samples. The area under the ROC curve, also known as ROC/AUC value, is a precision estimate expressed as a continuous value within a 0 to 1 range. The higher the ROC/AUC value, the more precise the logistic prediction model. The ROC/AUC value and the area under the ROC curve are among the most objective methods for the evaluation of binary classifiers (27, 31) and have previously been used to predict enterovirus presence based upon somatic coliphage concentrations in recreational waters (47). Multiple logistic regression and ROC curve analyses (27, 47) were executed using STATA software version 13 (48).

The recommended cut-off points for ROC/AUC were used in this study to determine the logistic regression model's discrimination ability: 0 to 0.5, null discrimination; 0.7 to 0.8, acceptable discrimination; 0.8 to 0.9, excellent discrimination; and 0.9 - 1.0, exceptional discrimination (31, 49). The multiple logistic regression model's discrimination ability must be at least acceptable (ROC/AUC ≥ 0.7) in order to identify a statistically-sound WWTP effluent indicator threshold concentration associated with an increased probability of human enteric

308 virus detection. Additionally, the indicator concentration parameter in the multiple logistic
309 regression model must have a significant association (p-value <0.05) with the detection of any
310 human enteric virus. The WWTP effluent somatic coliphage multiple regression model was
311 the only model to comply with the aforementioned criteria. The somatic coliphage threshold
312 concentration was identified at the concentration associated with the greatest sensitivity and
313 specificity in the ROC analysis.

314 The calculated somatic coliphage threshold concentration was evaluated for its ability
315 to identify human enteric virus PCR-positive WWTP effluent samples (27, 31). True-positive
316 (*i.e.*, PCR-positive for any of the human enteric viruses analyzed and somatic coliphage
317 concentration equal to or above the threshold), true-negative (*i.e.*, PCR-negative for any of the
318 human enteric viruses analyzed and somatic coliphage concentration below the threshold),
319 false-positive (*i.e.*, PCR-negative for any of the human enteric viruses analyzed and somatic
320 coliphage concentration equal to or above the threshold), and false-negative (*i.e.*, PCR-positive
321 for any of the human enteric viruses analyzed and somatic coliphage concentration below the
322 threshold) samples were calculated. Finally, the positive predictive value (*i.e.*, probability of
323 being PCR-positive for any of the human enteric viruses analyzed and the sample exceeded the
324 indicator threshold) and the negative predictive value (*i.e.*, probability of being PCR-negative
325 for any of the human enteric viruses analyzed and the sample was below the indicator
326 threshold) were calculated.

327 ***2.7 Quantitative microbial risk assessment for indirect reuse of wastewater treatment plant***
328 ***effluent meeting the somatic coliphage threshold***

329 In order to understand the health risks associated with the proposed somatic coliphage
330 threshold, quantitative microbial risk assessment was executed for a hypothetical wastewater

reuse scenario using EV, HAV, and NoVGI as reference pathogens. RV was not included in this analysis because adults are not susceptible to RV (50). NoVGII was not included because no dose-response curve currently exists (62). The annual disease burden for an adult farmer indirectly irrigating with WWTP effluent meeting the somatic coliphage threshold was estimated in 'R' Version 3.5.3 (www.rproject.org). For each model parameter defined as a distribution, a set of 10,000 random values was used to calculate the annual disease burdens in order to account for the uncertainty and variability associated with the model parameters.

First, daily exposure was defined for an adult farmer indirectly using the WWTP effluents from this study to irrigate crops, using the following equation (Eq. 3) for each enteric virus and parameter values/distributions (Table 2):

$$dose = v \times \left(\frac{c \times e^{-k_d t}}{(1+d)} \right)$$

where c is the WWTP effluent virus concentration when somatic coliphage concentrations are below the threshold, v is the volume of water accidentally ingested by the adult farmer irrigating on one day, d is the dilution factor from the WWTP effluent mixing in the river, k_d is the mean virus decay rate constant, and t is decay time (*i.e.*, the time the virus was in the river prior to irrigation). Similar to other studies, it was assumed that 1 mL of water was accidentally ingested per day of exposure (51, 52). The infectious enteric virus concentration in the WWTP effluent was defined as a uniform distribution between 0 and the maximum theoretical process limit of detection. Virus decay followed a first-order decay equation (53), using mean decay rates determined from experiments with similar conditions to those in the Virilla River (54–56). The WWTP effluent in this study is indirectly reused at distances ranging from 1 m to 3 km from WWTP discharge; thus, decay time was defined as a uniform distribution between 0 and 1 days (57). Since the dilution factor can vary greatly over time and

by season, the dilution factor was defined as a uniform distribution between a conservative dilution factor (99:1) and a maximum dilution factor (50,000:1) (53).

The daily probability of infection (P_{inf}) for each virus was then calculated using the dose previously calculated (Eq. 3) and the previously published dose-response curves and parameters distributions (Table 2). Briefly, the exponential dose-response curve was used for EV, which was derived from a study with pigs and porcine enterovirus type 7 (58). The exponential dose-response curve was also used for HAV, derived from a HAV human challenge study (59). For NoVGI, the fractional Poisson dose-response curve, derived from NoVGI human challenge studies, was used (60). Since there is no agreement among the scientific community with respect to NoV dose response parameters, they were described as recommended (62). The NoVGI aggregation factor (μ) was described as distribution ranging from minimum to maximum aggregation. The NoVGI genetically susceptible fraction of the population (*p*) was adjusted to represent Costa Rica's demographics (61, 62).

Subsequently, the daily probability of illness (*Pill*) for each virus was calculated with the following equation (Eq. 4):

$$P_{ill} = P_{inf} \times M$$

where P_{inf} is the probability of infection previously calculated and M is the morbidity ratio (Table 2 (50, 51, 59, 63–66)). The annual risk of illness (*Pa*) for each virus was then calculated as follows (Eq. 5):

$$P_a = 1 - (1 - P_{ill})^n$$

where P_{ill} is the daily risk of illness (Eq. 4) and n is the number of days adult farmers are exposed each year. Similar to other wastewater reuse irrigation studies, it was assumed that farmers irrigated 75 days per year (51, 65, 66). Finally, the annual disease burden (*DB*; daily

374 life adjusted years (DALYS)/person) for each virus was estimated as follows (Eq. 6):

$$DB = P_a \times B \times S_f$$

375 where *Pa* is the annual risk of illness (Eq. 5), *B* is the disease burden per case of illness, and S_f
376 is the susceptible fraction of the population (Table 2). The disease burden per case of illness
377 (*B*) was not available for Costa Rica (middle-income country); thus, it was defined as a
378 uniform distribution with minimum and maximum values identified for developing and
379 developed countries (50, 67–69). The NoV susceptible fraction of the population (S_f) was
380 defined as a uniform distribution for the Costa Rican demographic (61, 62). The EV
381 susceptible fraction of the population (S_f) was assumed to be 1 given high EV diversity (51).
382 The HAV susceptible fraction of the population was 0.717, as defined by seroprevalence in
383 adult Costa Rican population (28). Finally, the cumulative annual disease burden per person
384 from the three reference viruses was calculated by adding together the annual disease burden
385 (*DB*) for each virus. Since the dose calculation usually has the most significant influence on
386 model outputs (70) this quantitative microbial risk assessment's sensitivity to the exposure
387 assessment (Eq. 3) input parameters was tested by calculating the Spearman rank order
388 coefficients between the simulated input parameters and the estimated cumulative annual
389 disease burden ($\alpha = 0.05$).
390
391 **3. Results and Discussion**
392
393 ***3.1 Fecal coliforms and somatic coliphage in untreated wastewater***
394 For the five WWTPs investigated during this study, the mean (+/- standard deviation)
395 fecal coliform influent concentration was estimated as 6.8- ± 6.8-$\log_{10}$ MPN/ 100 mL, similar to

those summarized in the literature (71). The mean (+/- standard deviation) somatic coliphage influent concentration was 8.7- ± 9.0-$\log_{10}$ PFU/ 100 mL, which is 3-$\log_{10}$ PFU/ 100 mL greater than the mean concentration calculated in a recent global meta-analysis (Figure 2) (11). It is important to note that this recent meta-analysis did not include any Latin American countries and identified statistically significant differences between the geographical locations studied (11). The mean somatic coliphage influent concentrations reported in this study are more comparable to those in Argentina and Colombia, which are likely more similar to those in Costa Rica due to geographic location and water usage (72).

3.2 Fecal coliforms and somatic coliphage highly variable in treated wastewater

Both fecal coliforms and somatic coliphage concentrations were highly variable in the WWTP effluent studied, with mean and standard deviations estimated as 6.1- ± 6.6-$\log_{10}$ MPN/ 100 mL and 3.2- ± 3.1-$\log_{10}$ PFU/ 100 mL, respectively (Figure 2). The effluent fecal coliforms and somatic coliphage concentrations were similar to other WWTP studies (3, 20, 73, 74). Variability in the WWTP operational conditions (e.g., concentration of mixed liquor suspended solids, temperature, and biochemical oxygen demand (BOD)) are likely responsible for the indicator variability observed in this study (10, 72, 75, 76). Globally, fecal coliforms and somatic coliphage mean concentrations were significantly lower in the effluent in comparison to the influent ($p<0.0001$). However, with respect to WWTPs individually, mean fecal coliforms and somatic coliphage concentrations were lower in effluents than in influents at three of the five WWTPs ($p = 6.0\times10^{-6}$ to 0.02) and all five of the WWTPs ($p = 6.0\times10^{-7}$ to 0.04), respectively. Globally, the fecal coliforms and somatic coliphage mean (+/- standard deviation) log reduction

values were 0.99- ± 1.33-$\log_{10}$ and 2.70- 2.60-$\log_{10}$, respectively, and coincided with ranges previously reported (1).

3.3 Human enteric viruses frequently detected in (un)treated wastewater

Human enteric viruses (EV, HAV, NoVGI, NoVGII, or RV) were detected in WWTP influent and effluent at variable frequencies (Table 3). No statistical difference with respect to the frequency of human enteric virus detection was found between the WWTP influent and effluent ($p > 0.45$). RV was the most frequently detected in both influent and effluent samples (47% and 39%, respectively; Table 3), followed by NoV (GI and GII; 39% and 36%, respectively). Globally, NoVGI was detected two times more frequently than NoVGII. Less than 25% of the samples were positive for EV and less than 10% of the samples for HAV. It is important to mention that EV and HAV were analyzed in 117 out of 119 water samples; meanwhile, RV and NoV were analyzed in two-thirds of the samples (n = 79 and 80, respectively). Similar to all PCR-based analyses, it is possible that samples with undetected viruses had virus concentrations below the method detection limits or that inhibitors decreased RT-PCR efficiency (24). Additionally, a mixture of end-point and qPCR assays were effectively used in this study because improved resources were not logistically available. The lack of available resources is common in middle- and low-income countries because funds are limited, and supplies are often more expensive as well as difficult to import. When possible, future studies should use just one type of PCR-based analysis.

The RV and NoV data presented in this study corroborate with RV and NoV epidemiologies in Central and South American countries, in which they are present throughout the year and peak during the dry season (December – May) (77–79). Similar to our study,

441 NoVGI has been previously quantified in Costa Rican wastewater year-round, with peaks in the
442 dry season, at a WWTP in the Province of Puntarenas; in contrast, RV was the most frequently
443 detected virus in our study and the lowest quantified in Symonds *et al*. (80). The difference in
444 RV prevalence between the two studies is likely due to RV epidemiology in Costa Rica, where
445 RV infection is more frequent in the Greater Metropolitan Area compared to coastal regions
446 (such as the WWTP in Puntarenas) (81). With respect to EV, detection was very low (22%) in
447 influent and effluent in comparison to the USA (e.g., > 92% (25)). This difference may be due to
448 differences in epidemiology and/or methods between the two studies; however, it is difficult to
449 ascertain the origin of these differences because Central American EV epidemiology data is
450 limited.
451
452 ***3.3 Fecal coliforms do not correlate with human enteric virus detection and no threshold***
453 ***identified***
454 Multiple logistic regression models were used to analyze the statistical relationship
455 between fecal coliform concentrations and the detection of human enteric viruses at influent
456 and effluent wastewater samples. The estimated parameters from this logistic regression model
457 were -3.47×10^{-07} ($p = 0.258$) for fecal coliforms concentrations and 0.8881 ($p = 0.297$) for
458 dry/rainy season. According to the model, fecal coliforms do not correlate with human enteric
459 virus detection in WWTP influent or effluent (OR = 0.99, $p = 0.26$). Despite the lack of
460 relationship between fecal coliforms and human enteric viruses detection, ROC analysis was
461 used to estimate a possible fecal coliform concentration associated with the detection of any of
462 the five human enteric viruses analyzed (*i.e.*, to identify an appropriate maximum fecal
463 coliform concentration associated with increased human enteric virus detection). The ROC

464 analysis for fecal coliforms and human enteric virus detection did not have acceptable
465 precision (ROC/AUC= 0.64). These findings corroborate with previous studies that did not
466 identify correlations between fecal coliform concentrations and human enteric virus detection
467 (10, 20, 41, 82).
468
469 ***3.4 Somatic coliphage correlate with human enteric virus detection and a threshold was***
470 ***identified***
471 Multiple logistic regression models were also used to analyze the statistical relationship
472 between somatic coliphage concentrations (PFU/ 100 mL) and human enteric virus detection
473 in WWTP influent and effluents. For the WWTP influent, the estimated multiple logistic
474 regression model parameters were 3.98×10^{-10} ($p = 0.074$) for somatic coliphage concentration
475 and 0.5606 ($p = 0.035$) for dry/rainy season. WWTP influent had a 75% probability of being
476 positive for at least one of the human enteric viruses studied during the dry season in
477 comparison to the rainy season (OR = 1.75, $p = 0.035$). Additionally, a significant correlation
478 between somatic coliphage concentrations and human enteric virus detection was identified in
479 WWTP effluent (OR = 1.00, $p = 0.01$), which was similar to those previously described by (3,
480 10, 83). For the WWTP effluent, the estimated multiple logistic regression model parameters
481 were -0.0004 ($p = 0.006$) for somatic coliphage concentration and 0.8881 ($p = 0.297$) for
482 dry/rainy season. It is important to note that season was not a significant predictor of human
483 enteric virus detection in WWTP effluent (OR = 2.43, $p = 0.297$).
484 In order to determine an appropriate somatic coliphage concentration associated with
485 an increased probability of human enteric virus detection, ROC analysis was used to estimate
486 the somatic coliphage concentration associated with human enteric virus detection (*i.e.*, any of

the five viruses) in WWTP effluent. The area under the ROC curve (AUC) was 0.7 (Figure 3); thus, it had an acceptable discrimination ability. The sensitivity and specificity curves intersect near the 0.526 probability cutoff, where the highest specificity (75%) and sensibility (54%) were found when the somatic coliphage concentration was 3.5-$\log_{10}$ PFU/ 100 mL ($p = 0.526$). Thus, this somatic coliphage threshold (3.5-$\log_{10}$ PFU/100 mL) was the concentration most likely associated with a lack of human enteric virus detection.

This somatic coliphage threshold was evaluated for its ability to identify human enteric virus PCR-positive WWTP effluent samples by calculating the Positive and Negative Predictive Values (31). The frequencies of true-/false-positives and true-/false-negatives were calculated for each enteric virus type (Table 4), which were used to calculate the Positive and Negative Predictive values. Positive Predictive Value was 46%; therefore, 46% of samples had somatic coliphage concentrations above the threshold and human enteric viruses were detected. The Negative Predictive Value was 33%; thus, 33% of samples had somatic coliphage concentrations less than the threshold and no human enteric viruses were detected. Using this threshold, only a 34.5% of the samples were classified as false-negative and would represent a possible human health risk (Table 4). Overall, 65.6% of WWTP effluent samples were safely classified using the proposed somatic coliphage threshold.

Similar to this study, low or undetectable human enteric virus concentrations were measured in WWTP effluent when somatic coliphage concentrations were below 3.5-$\log_{10}$ PFU/ 100 mL (3, 20, 73, 74, 83, 84). Nevertheless, it is important to note that the results of this study are directly dependent on the efficiencies and detection limits of the methods used. It is possible that the somatic coliphage threshold would be different if different methods (*e.g.*, virus concentration, RNA extraction) were used or if additional/fewer human enteric viruses had been

analyzed. Furthermore, this study does not take into consideration the detection of infectious human enteric viruses. Future studies are needed to explore and confirm the somatic coliphage threshold identified in this study. Specifically, studies are needed that take into consideration human enteric virus infectivity. It is also important to analyze how the use of different methods may or may not influence the somatic coliphage threshold identified. Interestingly, the somatic coliphage threshold identified in this study is similar to the threshold previously proposed two decades ago (3-$\log_{10}$ PFU/ 100 mL), even though this study were executed using different statistical and virus methods, only analyzed human EV, and used cell-culture methods (20).

3.5 Annual disease burden for indirectly reusing wastewater effluent below the proposed threshold

Quantitative microbial risk assessment was used to estimate the EV, HAV, NoVGI, as well as cumulative (all three viruses) annual disease burden for an adult farmer irrigating indirectly (75 days per year) with WWTP effluent below the proposed somatic coliphage threshold. The median cumulative annual disease burden per adult farmer was 2.52×10^{-5} DALYs (Figure 4), which is less than the recommendation of 10^{-4} (11, 85). EV contributed the most to the cumulative annual disease burden, followed by HAV and NoVGI. The exposure assessment parameter sensitivity analysis indicated that the daily volume ingested ($\rho = 0.479$, $p = 2.2 \times 10^{-16}$), WWTP effluent infectious enteric virus concentrations ($\rho = 0.471$, $p = 2.2 \times 10^{-16}$), and the dilution factor ($\rho = -0.466$, $p = 2.2 \times 10^{-16}$) were most influential on the cumulative annual disease burden estimates in comparison to the decay-related variables ($0.102 \leq | \rho | \leq 0.231$; $p < 2.2 \times 10^{-16}$). Furthermore, the NoVGI decay rate did not significantly correlate with the cumulative annual disease burden ($\rho = -0.017$, $p = 0.087$).

533 Based upon the sensitivity analysis results, it is likely that the cumulative annual disease
534 burdens may increase or decrease markedly if the estimated daily volume ingested and/or
535 WWTP effluent infectious enteric virus concentrations were higher or lower, respectively. In
536 order to incorporate uncertainty and variability in this study, WWTP effluent infectious enteric
537 virus concentrations were defined as a uniform distribution between 0 and the maximum
538 theoretical process limit of detection. It was assumed that all viruses were infectious; thus, it is
539 possible that the cumulative disease burden calculated overestimated risk. Additionally, the
540 cumulative disease burden calculated could underestimate the actual risk if the theoretical
541 process limit of detection was greater than the maximum value estimated. Nevertheless, the
542 theoretical process limit of detection took into account losses associated with virus concentration
543 and detection.

544 Since Costa Rican culture lacks habits associated with additional hand-mouth contact
545 (*e.g.*, Bolivia, chewing coca leaves), a point value traditionally used in quantitative microbial risk
546 assessment was used even though it can impact model output (51). Similarly, it was difficult to
547 identify the dilution factor of the WWTP effluent entering the river due to constant fluctuations
548 in river flow rates and volume. Consequently, this parameter was defined as a distribution
549 between a conservative (99:1) and maximum (50,000:1) assumption (53). If the dilution factor
550 was greater, then the cumulative annual disease burden estimates would be much lower. Finally,
551 cumulative annual disease burden would be greatly affected if the number of days farmers
552 irrigated indirectly with WWTP effluent were greater or less than the assumed 75 days.

553 While quantitative microbial risk assessment is a useful tool, it is based upon assumptions
554 that may or may not reflect reality. Consequently, this quantitative microbial risk assessment
555 incorporated the use of parameter distributions to account for this uncertainty and variability.

556 Nevertheless, it is important to recognize that the dose-response curves and parameter
557 distributions may not reflect realistic human populations, which can highly influence model
558 outputs (53, 70). This is particularly true for EV, which is based upon a non-human model (51,
559 58), and NoVGI because there is no agreement on which parameters are most appropriate (62).
560 Point values were also used for certain parameters when preferred values were previously
561 identified in other studies. Additionally, data from other contexts were used when contextualized
562 data were lacking. Similar to other studies, it is likely that these assumptions influenced the
563 model output (51,53, 70). Although wastewater contains a wide-variety of disease-causing
564 viruses, it is not possible to estimate the disease burden of all pathogens given the lack of
565 disease-related and dose-response data. Consequently, this study incorporated three reference
566 pathogens to calculate the cumulative annual disease burden associated with indirect wastewater
567 reuse with WWTP effluent meeting the somatic coliphage threshold.
568
569 ***3.6 Implications for activated sludge WWTP effluent management***
570 As far as we know, this is the first report of a statistically sound somatic coliphage
571 threshold estimation for WWTP effluent management. The use of a somatic coliphage threshold
572 of 3.5-$\log_{10}$ PFU/ 100 mL is an affordable alternative and/or complement to the virus log
573 reduction value multiple barrier system approach, and if implemented, could improve WWTP
574 effluent management in resource-limited regions that have been resistant to the aforementioned
575 approach. Thus, compliance with this threshold would assure lower enteric virus concentrations
576 discharged into nearby rivers with downstream uses in agriculture.
577 Additionally, the indirect reuse of WWTP effluent meeting the proposed somatic
578 coliphage threshold was associated with a median cumulative annual disease burden that

complies with the WHO recommendation (13, 85). Given the potential of the proposed somatic coliphage to improve activated sludge WWTP effluent management, further research is warranted to validate, improve, and optimize this threshold for use in Costa Rica. Future investigations should include improved disease burden estimates that contain the most context-appropriate data possible, especially with respect to the exposure assessment parameters. Additional research is also needed to validate the way in which such a threshold should be implemented (*e.g.*, geometric mean, single measurement, 95% percentile) to ensure improved wastewater effluent management, and ultimately better protect public health. Finally, the statistical approach presented here can be implemented in other regions to determine a logical and feasible metric to improve upon existing WWTP discharge legislation.

Acknowledgment

We would like to recognize the Vice Rector of Investigation at the University of Costa Rica (Vicerrectoría de Investigación de la Universidad de Costa Rica) for funding our research. Special thanks to the participants of the INISA-CDRC University of Miami Scientific Publications Workshop for their valuable suggestions and Lic. Eric Morales Mora from INISA-UCR for his help with creating Figure 1. E.M.S was funded by the US National Science Foundation grant OCE-1566562. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the US National Science Foundation. Finally, we would like to thank Dr. Walter Betancourt for his suggestions with respect to the manuscript.

601 **References**

602

603 1. Naughton C, Rousselot O. 2017. Activated Sludge, p. . *In* Mihelcic, J, Verbyla, M
604 (eds.), Global Water Pathogen Project, 1st ed. Michigan State University, Michigan.
605 2. De Luca G, Sacchetti R, Leoni E, Zanetti F. 2013. Removal of indicator bacteriophages
606 from municipal wastewater by a full-scale membrane bioreactor and a conventional
607 activated sludge process: Implications to water reuse. Bioresour Technol 129:526–531.
608 3. Dias E, Ebdon J, Taylor H. 2018. The application of bacteriophages as novel indicators
609 of viral pathogens in wastewater treatment systems. Water Res 129:172–179.
610 4. Gerba CP, Betancourt WQ, Kitajima M. 2017. How much reduction of virus is needed
611 for recycled water: A continuous changing need for assessment? Water Res 108:25–31.
612 5. Hewitt J, Leonard M, Greening GE, Lewis GD. 2011. Influence of wastewater treatment
613 process and the population size on human virus profiles in wastewater. Water Res
614 45:6267–6276.
615 6. Ruiz Fallas F. 2012. “Gestión de las Excretas y Aguas Residuales en Costa Rica”
616 Situación Actual y Perspectiva. San José.
617 7. Sano D, Amarasiri M, Hata A, Watanabe T, Katayama H. 2016. Risk management of
618 viral infectious diseases in wastewater reclamation and reuse: Review. Environ Int
619 91:220–229.
620 8. Sidhu JPS, Sena K, Hodgers L, Palmer A, Toze S. 2018. Comparative enteric viruses
621 and coliphage removal during wastewater treatment processes in a sub-tropical
622 environment. Sci Total Environ 616–617:669–677.
623 9. U.S. Environmental Protection Agency. 2012. Guidelines for Water Reuse. US Environ

624 Prot Agency Natl Risk Manag Res Lab US Agency Int Dev. Washington, DC, United
625 States of America.
626 10. McMinn BR, Ashbolt NJ, Korajkic A. 2017. Bacteriophages as indicators of fecal
627 pollution and enteric virus removal. Lett Appl Microbiol 65:11–26.
628 11. Nappier SP, Hong T, Ichida A, Goldstone A, Eftim SE. 2019. Occurrence of coliphage
629 in raw wastewater and in ambient water: A meta-analysis. Water Res 153:263–273.
630 12. National Research Council. 2011. Water Reuse : Expanding the Nation's Water Supply
631 Through Reuse of Municipal Wastewater Understanding the Risks, 1st ed. The National
632 Academies Press, Washington, DC.
633 13. World Health Organization. 2006. WHO Guidelines for the safe use of wastewater,
634 excreta and greywater in Agriculture. WHO Libr Cat Data. World Health Organization,
635 Geneva.
636 14. Committee on Indicators for Waterborne Pathogens, Board on Life Sciences, Water
637 Science and Technology Board, Division on Earth and Life Studies, National Research
638 Council. 2004. Indicators for Waterborne PathogensIndicators for Waterborne
639 Pathogens. National Academies Press, Washington, D.C.
640 15. Baggi F, Demarta A, Peduzzi R. 2001. Persistence of viral pathogens and
641 bacteriophages during sewage treatment: lack of correlation with indicator bacteria. Res
642 Microbiol 152:743–751.
643 16. Carducci A, Battistini R, Rovini E, Verani M. 2009. Viral Removal by Wastewater
644 Treatment: Monitoring of Indicators and Pathogens. Food Environ Virol 1:85–91.
645 17. U.S. Environmental Protection Agency. 2015. Review of coliphages as possible
646 indicators of fecal contamination for ambient water quality. Washington, DC.

18. Savichtcheva O, Okabe S. 2006. Alternative indicators of fecal pollution: Relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. Water Res 40:2463–76.

19. Yates M V. 2007. Classical Indicators in the 21st Century-Far and Beyond the Coliform. Water Environ Res 79:279–286.

20. Gantzer C, Maul A, Audic JM, Schwartzbrod L. 1998. Detection of Infectious Enteroviruses, Enterovirus Genomes, Somatic Coliphages, and Bacteroides fragilis Phages in Treated Wastewater. Appl Environ Microbiol 64:4307–4312.

21. Pillai SD. 2006. Bacteriophages as Fecal Indicator Organisms, p. 205–222. *In* Goyal, SM (ed.), Viruses in Foods. Springer US, Boston, MA.

22. Ito T, Kato T, Hasegawa M, Katayama H, Ishii S, Okabe S, Sano D. 2016. Evaluation of virus reduction efficiency in wastewater treatment unit processes as a credit value in the multiplebarrier system for wastewater reclamation and reuse. J Water Health 14:879–889.

23. Gerba CP, Betancourt WQ, Kitajima M, Rock CM. 2018. Reducing uncertainty in estimating virus reduction by advanced water treatment processes. Water Res 133:282–288.

24. Haramoto E, Kitajima M, Hata A, Torrey JR, Masago Y, Sano D, Katayama H. 2018. A review on recent progress in the detection methods and prevalence of human enteric viruses in water. Water Res 135:168–186.

25. Kitajima M, Iker BC, Pepper IL, Gerba CP. 2014. Relative abundance and treatment reduction of viruses during wastewater treatment processes — Identification of potential viral indicators. Sci Total Environ 488–489:290–296.

670 26. Pouillot R, Van Doren JM, Woods J, Plante D, Smith M, Goblick G, Roberts C, Locas
671 A, Hajen W, Stobo J, White J, Holtzman J, Buenaventura E, Burkhardt W, Catford A,
672 Edwards R, DePaola A, Calci KR. 2015. Meta-Analysis of the Reduction of Norovirus
673 and Male-Specific Coliphage Concentrations in Wastewater Treatment Plants. Appl
674 Environ Microbiol 81:4669–4681.
675 27. Dorevitch S. 2016. Comments on the US EPA "Review of Coliphages as Possible
676 Indicators of Fecal Contamination for Ambient Water Quality." Chicago.
677 28. Taylor ML, García Z, Holst I, Somogye T, Cunningham L, Visoná KA. 2001.
678 Seroprevalencia de los virus de la Hepatitis A y B en grupos etarios de Costa Rica. Acta
679 Med Costarric 43:153–158.
680 29. World Health Organization. 2017. Guidelines for drinking-water quality: fourth edition
681 incorporating the first addendum, 4th ed. Geneva.
682 30. White SE. 2020. Statistical Methods for Multiple VariablesBasic & Clinical
683 Biostatistics, 5th ed. McGraw-Hill Education, New York, NY.
684 31. White SE. 2020. Methods of Evidence-Based Medicine and Decision AnalysisBasic &
685 Clinical Biostatistics, 5th ed. McGraw-Hill Education, New York, NY.
686 32. Mora-Alvarado DA, Portuguez-Barquero CF. 2016. Cobertura de la disposición de
687 excretas en Costa Rica en el periodo 2000-2014 y expectativas para el 2021. Rev Tecnol
688 en Marcha 29:43.
689 33. APHA (American Public Health Asociation), AWWA (American Water Works
690 Association), WEF (Water Environment Federation). 2005. Standard Methods for the
691 examination of water and wastewater, 21st ed. American Public Health Association,
692 Washington, DC.

693 34. Deshpande JM, Shetty SJ, Siddiqui ZA. 2003. Environmental Surveillance System To
694 Track Wild Poliovirus Transmission. Appl Environ Microbiol 69:2919–2927.
695 35. Cashdollar JL, Wymer L. 2013. Methods for primary concentration of viruses from
696 water samples: a review and meta-analysis of recent studies. J Appl Microbiol 115:1–
697 11.
698 36. Dulbecco R, Vogt M. 1954. Plaque Formation and Isolation of Pure Lines with
699 Poliomyelitis Viruses. J Exp Med 99:167–182.
700 37. Sarmiento L, Mas P, Goyenechea A, Palomera R, Morier L, Capó V, Quintana I, Santin
701 M. 2001. First Epidemic of Echovirus 16 Meningitis in Cuba. Emerg Infect Dis 7:887–
702 889.
703 38. Arauz-Ruiz P, Sundqvist L, García Z, Taylor L, Visoná K, Norder H, Magnius LO.
704 2001. Presumed Common Source Outbreaks of Hepatitis A in an Endemic Area
705 Confirmed by Limited Sequencing Within the VP1 Region. J Med Virol 65:449–456.
706 39. Pang X, Cao M, Zhang M, Lee B. 2011. Increased sensitivity for various rotavirus
707 genotypes in stool specimens by amending three mismatched nucleotides in the forward
708 primer of a real-time RT-PCR assay. J Virol Methods 172:85–87.
709 40. Japhet MO, Adesina OA, Famurewa O, Svensson L, Nordgren J. 2012. Molecular
710 epidemiology of rotavirus and norovirus in Ile-Ife, Nigeria: High prevalence of G12P[8]
711 rotavirus strains and detection of a rare norovirus genotype. J Med Virol 84:1489–1496.
712 41. Harwood VJ, Levine AD, Scott TM, Chivukula V, Lukasik J, Farrah SR, Rose JB.
713 2005. Validity of the Indicator Organism Paradigm for Pathogen Reduction in
714 Reclaimed Water and Public Health Protection. Appl Environ Microbiol 71:3163–3170.
715 42. Miranda JA, Steward GF. 2017. Variables influencing the efficiency and interpretation

of reverse transcription quantitative PCR (RT-qPCR): An empirical study using Bacteriophage MS2. J Virol Methods 241:1–10.

43. Macherey-Nagel. 2014. Viral RNA and DNA isolation User manual: NucleoSpin® RNA Virus/ NucleoSpin® RNA Virus F. Düren.

44. Solano Barquero M, Chacón Jiménez L, Barrantes Jiménez K, Achí Araya R. 2012. Implementación de dos métodos de recuento en placa para la detección de colifagos somáticos, aportes a las metodologías estándar. Rev Peru Biol 19:335–340.

45. Lee L. 2017. NADA: Nondetects and Data Analysis for Environmental Data. R package version 1.6-1.

46. McDonnell G, Russell AD. 1999. Antiseptics and disinfectants: Activity, action, and resistance. Clin Microbiol Rev 12:147.

47. Mocé-Llivina L, Lucena F, Jofre J. 2005. Enteroviruses and Bacteriophages in Bathing Waters. Appl Environ Microbiol 71:6838–6844.

48. StataCorp. 2013. Stata Statistical Software: Release 13. 13. StataCorp LP, Texas.

49. Berrar D, Flach P. 2011. Caveats and pitfalls of ROC analysis in clinical microarray research (and how to avoid them). Brief Bioinform 13:83–97.

50. Havelaar AH, Melse JM. 2003. Quantifying public health risk in the WHO Guidelines for Drinking-Water QualityRIVM report 734301022/2003. Bilthoven.

51. Symonds EM, Verbyla ME, Lukasik JO, Kafle RC, Breitbart M, Mihelcic JR. 2014. A case study of enteric virus removal and insights into the associated risk of water reuse for two wastewater treatment pond systems in Bolivia. Water Res 65:257–270.

52. Ottoson J, Stenström TA. 2003. Faecal contamination of greywater and associated microbial risks. Water Res 37:645–655.

739 53. Sunger N, Hamilton KA, Morgan PM, Haas CN. 2019. Comparison of pathogen-
740 derived 'total risk' with indicator-based correlations for recreational (swimming)
741 exposure. Environ Sci Pollut Res 26:30614–30624.
742 54. Boehm AB, Silverman AI, Schriewer A, Goodwin K. 2019. Systematic review and
743 meta-analysis of decay rates of waterborne mammalian viruses and coliphages in
744 surface waters. Water Res 164:114898.
745 55. De Roda Husman AM, Lodder WJ, Rutjes SA, Schijven JF, Teunis PFM. 2009. Long-
746 term inactivation study of three enteroviruses in artificial surface and groundwaters,
747 using PCR and cell culture. Appl Environ Microbiol 75:1050–1057.
748 56. Bae J, Schwab KJ. 2008. Evaluation of murine norovirus, feline calicivirus, poliovirus,
749 and MS2 as surrogates for human norovirus in a model of viral persistence in surface
750 water and groundwater. Appl Environ Microbiol 74:477–484.
751 57. Quesada Q. ME, Waylen PR. 2012. Diferencias hidrológicas anuales y estacionales en
752 regiones adyacentes: estudio de las subcuencas de los ríos Virilla y Grande de San
753 Ramón, Costa Rica. Cuad Geogr Rev Colomb Geogr 21:167–179.
754 58. Cliver DO. 1981. Experimental Infection by Waterborne Enteroviruses. J Food Prot
755 44:861–865.
756 59. Weir MH. 2019. A data simulation method to optimize a mechanistic dose-response
757 model for viral loads of hepatitis A. Microb Risk Anal 100102.
758 60. Messner MJ, Berger P, Nappier SP. 2014. Fractional poisson - A simple dose-response
759 model for human norovirus. Risk Anal 34:1820–1829.
760 61. Morera B, Barrantes R, Marin-Rojas R. 2003. Gene admixture in the Costa Rican
761 population. Ann Hum Genet 67:71–80.

762 62. Van Abel N, Schoen ME, Kissel JC, Meschke JS. 2017. Comparison of Risk Predicted
763 by Multiple Norovirus Dose-Response Models and Implications for Quantitative
764 Microbial Risk Assessment. Risk Anal 37:245–264.
765 63. Ward R, Krugman S, Giles JP, Jacobs AM, Bodansky O. 1958. Infectious Hepatitis. N
766 Engl J Med 258:407–416.
767 64. Soller JA, Schoen M, Steele JA, Griffith JF, Schiff KC. 2017. Incidence of
768 gastrointestinal illness following wet weather recreational exposures: Harmonization of
769 quantitative microbial risk assessment with an epidemiologic investigation of surfers.
770 Water Res 121:280–289.
771 65. Seidu R, Heistad A, Amoah P, Drechsel P, Jenssen PD, Stenström TA. 2008.
772 Quantification of the health risk associated with wastewater reuse in Accra, Ghana: A
773 contribution toward local guidelines. J Water Health 6:461–471.
774 66. Mara DD, Sleigh PA, Blumenthal UJ, Carr RM. 2007. Health risks in wastewater
775 irrigation: Comparing estimates from quantitative microbial risk analyses and
776 epidemiological studies. J Water Health 5:39–50.
777 67. Mok HF, Barker SF, Hamilton AJ. 2014. A probabilistic quantitative microbial risk
778 assessment model of norovirus disease burden from wastewater irrigation of vegetables
779 in Shepparton, Australia. Water Res 54:347–362.
780 68. Howard G, Feroze Ahmed M, Gaifur Mahmud S, Teunis P, Davison A, Deere D. 2007.
781 Disease burden estimation to support policy decision-making and research prioritization
782 for arsenic mitigation. J Water Health 5:67–81.
783 69. Prüss-Üstün A, Bos R, Gore F, Bartram J. 2008. Safer water, better health: costs,
784 benefits and sustainability of interentions to protect and promote healthWorld Health

Organization.

70. Verbyla ME, Symonds EM, Kafle RC, Cairns MR, Iriarte M, Mercado Guzmán A, Coronado O, Breitbart M, Ledo C, Mihelcic JR. 2016. Managing Microbial Risks from Indirect Wastewater Reuse for Irrigation in Urbanizing Watersheds. Environ Sci Technol 50:6803–6813.

71. Harwood V, Shanks O, Korajkic A, Verbyla M, Ahmed W, Iriarte M. 2019. General and host-associated bacterial indicators of faecal pollution, p. . *In* Farnleitner, A, Blanch, A (eds.), Global Water Pathogen Project. Michigan State University, Michigan.

72. Lucena F, Duran AE, Moron A, Calderon E, Campos C, Gantzer C, Skraber S, Jofre J. 2004. Reduction of bacterial indicators and bacteriophages infecting faecal bacteria in primary and secondary wastewater treatments. J Appl Microbiol 97:1069–1076.

73. Petrinca AR, Donia D, Pierangeli A, Gabrieli R, Degener AM, Bonanni E, Diaco L, Cecchini G, Anastasi P, Divizia M. 2009. Presence and environmental circulation of enteric viruses in three different wastewater treatment plants. J Appl Microbiol 106:1608–1617.

74. Francy DS, Stelzer EA, Bushon RN, Brady AMG, Williston AG, Riddell KR, Borchardt MA, Spencer SK, Gellner TM. 2012. Comparative effectiveness of membrane bioreactors, conventional secondary treatment, and chlorine and UV disinfection to remove microorganisms from municipal wastewaters. Water Res 46:4164–4178.

75. Gerba CP, Kitajima M, Iker B. 2013. Viral presence in waste water and sewage and control methods, p. 293–315. *In* Viruses in Food and WaterHolden, Ne. Elsevier, Oxford.

76. Zhang C-M, Xu L-M, Xu P-C, Wang XC. 2016. Elimination of viruses from domestic

wastewater: requirements and technologies. World J Microbiol Biotechnol 32:69.

77. Bucardo F, Lindgren P, Svensson L, Nordgren J. 2011. Low Prevalence of Rotavirus and High Prevalence of Norovirus in Hospital and Community Wastewater after Introduction of Rotavirus Vaccine in Nicaragua. PLoS One 6:e25962.

78. da Silva Poló T, Peiró JR, Mendes LCN, Ludwig LF, de Oliveira-Filho EF, Bucardo F, Huynen P, Melin P, Thiry E, Mauroy A. 2016. Human norovirus infection in Latin America. J Clin Virol 78:111–119.

79. Patel MM, Pitzer V, Alonso W, Vera D, Lopman B, Tate J, Viboud C, Parashar UD. 2013. Global Seasonality of Rotavirus Disease. Pediatr Infect Dis J 130:1514–1523.

80. Symonds EM, Young S, Verbyla ME, McQuaig-Ulrich SM, Ross E, Jiménez JA, Harwood VJ, Breitbart M. 2017. Microbial source tracking in shellfish harvesting waters in the Gulf of Nicoya, Costa Rica. Water Res 111:177–184.

81. Ministerio de Salud. 2019. Vigilancia de la Salud: Enfermedades de Notificación Individual.

82. Wu J, Long SC, Das D, Dorner SM. 2011. Are microbial indicators and pathogens correlated? A statistical analysis of 40 years of research. J Water Health 9:265–278.

83. Costán-Longares A, Montemayor M, Payán A, Méndez J, Jofre J, Mujeriego R, Lucena F. 2008. Microbial indicators and pathogens: Removal, relationships and predictive capabilities in water reclamation facilities. Water Res 42:4439–4448.

84. Gomila M, Solis JJ, David Z, Ramon C, Lalucat J. 2008. Comparative reductions of bacterial indicators, bacteriophage-infecting enteric bacteria and enteroviruses in wastewater tertiary treatments by lagooning and UV-radiation. Water Sci Technol 58:2223–2233.

831 85. Rose JB, Hofstra N, Murphy HM, Verbyla ME. 2019. What Is Safe Sanitation? J
832 Environ Eng 145:02519002.
833

Table 1. RT-PCR and RT-qPCR assays used to detect enterovirus, hepatitis A, rotavirus group A, and norovirus genotypes I and II.

Virus	Gene	Primer Sequence	PCR conditions	Primer Final concentration	Reference
EV	5' Non coding region	EV1: 5' ATTGTCACCATAAGCAGCCA 3' EV2: 5' TCCGGCCCCTGAATGCGGCTAATCC 3' EV3: 5' ACACGGACACCCAAAGTAGTCGGTTCC 3' EV4: 5' TCCGGCCCCTGAATGCGGCTAATCC3'	PCRI: EV1/EV2 Activation cycle (95 ºC 5 min) 30 cycles: 95 ºC 45s, 55 ºC 45 s, and 70 ºC 45s PCRII: EV3/EV4 Activation cycle (95 ºC 5 min) 30 cycles: 95 ºC 45s, 55 ºC 45 s, and 70 ºC 45s	EV1: 6.1 mM EV2: 7.6 mM EV3: 8.2 mM EV3: 7.6 mM	(37)
HAV	VP1 region	HA1: 5' TTGCTCCTCTTTATCATGCTATG 3' HA3: 5' TGGTTAAATCTAATGGTCCTCTATA 3'	Activation cycle (95 ºC 5 min) 40 cycles: 95 ºC 30s, 46 ºC 30 s, and 70 ºC 30s	HA1: 8.7 mM HA3: 9.6 mM	(38)

RV	NSP-3	ROTAS1: 5' ACCATCTTCACgTAACCCTC 3' ROTAS2: 5' ACCATCTACACATGACCCTC 3' ROTAA: 5' CACATAACGCCCCTATAGCC 3' ROTAP: [6FAM]- GGGGATGAGCACAATAGTTAAAAGCTAACA CTGTCAA -[BBQ]	Activation cycle (95 ºC 5 min) 40 cycles: 95 º 20 s, 60 ºC 40 s	ROTAS1: 0.2 pM ROTAS2: 0.2 pM ROTAA: 0.2 pM ROTAAP: 0.2 pM	(39)
NoV GI	ORF1	NVG1F: 5'CGYTGGATGCGNTTCCATGA 3' NVG1R: 5' GTCCTTAGACGCCATCATC 3' G1-prob: [6FAM]-AGATYGCGRTCYCCTGTCCA- [BHQ1]	Activation cycle (95 ºC 5 min) 40 cycles: 95 º 15 s, 56 ºC 60 s	NVG1F: 0.2 pM NVG1R: 0.2 pM G1-prob: 0.2 pM	(40)
NoV GII	ORF1	NVG2F: 5'ATGTTYAGRTGGATGAGRTTYTC 3' COG2R: 5' TCgACgCCATCTTCATTCACA 3' G2-prob:	Activation cycle (95 ºC 5 min) 40 cycles: 95 º 15 s, 56 ºC 60 s	NVG2F: 0.2 pM COG2R: 0.2 pM G2-prob: 0.2	(21)

[JOE]TGGGAGGGCGATCGCAATCT[BHQ1]	pM

Table 2. Quantitative microbial risk model parameter values/distributions, and dose-response equations.

Parameter	**Units**	**Value or distribution**	**Reference**
Virus concentration in WWTP effluent (c)	Virus/1 mL	uniform(0, 182)	This study.
Volume of water ingested (v)	mL	1	(51, 52)
Dilution factor (d)	proportion	uniform(99,50000)	(53)
Time in river (t)	day	uniform(0,1)	(53, 57)
Mean decay rate (k_d)	day^{-1}		
enterovirus		0.028	(55)
hepatitis A		0.22	(54)
norovirus genotype I		0.08	(56)
Dose-response			
enterovirus	Exponential	$P_{inf} = 1 - e^{-d \times k}$ k = uniform(0.00291, 0.00562)	(58)
hepatitis A	Exponential	$P_{inf} = 1 - e^{-d \times k}$ k = uniform(0.00005871, 0.001191)	(59)
norovirus genotype I	Fractional Poisson	$P_{inf} = P \times [1 - e^{\frac{-d}{\mu}}]$ P = uniform(0.87, 1) μ = uniform(1, 1106)	(60 - 62)
Morbidity Ratio (M)			

$M_{enterovirus}$	median	0.9	(50)
$M_{hepatitis\ A}$		uniform(0.25, 0.92)	(59, 63)
$M_{norovirus}$		uniform(0.3, 1)	(64)
Total annual days of irrigation	days/farmer	75	(51, 65-66)
Disease burden per illness (B)	DALYS/case of illness		
$B_{enterovirus}$		uniform(0.0024,0.0150)	(68, 69)
$B_{hepatitis\ A}$		uniform(0.0761, 0.191)	(50)
$B_{norovirus}$		uniform(0.000371, 0.00623)	(67)
Susceptible fraction of population (S)	proportion		
$S_{enterovirus}$		1	(51)
$S_{hepatitis\ A}$		0.717	(28)
$S_{norovirus}$		uniform(0.87,1.00)	(61, 62)

Table 3. Positive samples (%) for human enteric viruses in wastewater influent and effluent from five activated sludge wastewater treatment plants in Costa Rica, 2013.

Variable	Influent (No. positives / No. samples) (%)	Effluent (No. positives / No. samples) (%)	Total (No. positives / No. samples) (%)	p*
Any viral pathogen**	32/60 (53%)	31/59 (53%)	63/119 (53%)	0.93
Enterovirus	13/59 (22%)	13/58 (22%)	26/117 (22%)	0.96
Hepatitis A	5/59 (8%)	3/58 (5%)	8/117 (7%)	0.48
Rotavirus	18/38 (47%)	16/41 (39%)	34/79 (43%)	0.45
Norovirus GI	16/38 (39%)	13/34 (36%)	29/72 (37%)	0.74
Norovirus GII	9/38 (24%)	4/34 (12%)	13/72 (18%)	0.19
All Norovirus	16/41 (39%)	14/39 (36%)	30/80 (37%)	0.77

* Person Chi-square results for differences in detection of pathogens.

** Total number of water samples positive for any pathogenic virus.

Table 4. Relationship between human enteric virus detection and somatic coliphage concentrations above calculated threshold in wastewater treatment plant (WWTP) effluent.

	WWTP effluent samples with somatic coliphage detection above threshold			
Human Enteric Virus*	**True Positive (%)**	**False Positive (%)**	**False Negative (%)**	**True Negative (%)**
Enterovirus (n=58)	6 (10.3)	22 (37.9)	7 (12.1)	23 (39.7)
Hepatitis A Virus (n=58)	0 (0)	28 (48.3)	3 (5.2)	27 (46.6)
Rotavirus (n=41)	6 (14.6)	23 (56)	10 (24.4)	12 (29.3)
Norovirus (n=40)	6 (15.8)	13 (31.7)	7 (18.4)	14 (36.8)
Any virus (n=58)	13 (22.4)	15 (25.9)	20 (34.5)	10 (17.2)

* Some samples were positive for more than one human enteric virus

Figure Legends

Figure 1. San José Province depicting the location of the five wastewater treatment plants included in the study. The San José Province is located at an altitude of 760 – 1,230 m above sea level and has an average temperature of 22°C year-round. Annual average precipitation ranges from 2,000–3,000 mm.

Figure 2. Global somatic coliphage and fecal coliform concentrations at influent and effluent of WWTP by sampling period.

Figure 3. Area under Receiver Operating Characteristic (ROC) curve for the multiple logistic regression model of somatic coliphage concentrations as a function of human enteric virus detection in conventional activated sludge WWTP effluent.

Figure 4. The estimated annual disease burden for an adult farmer indirectly irrigating with wastewater treatment plant effluent below the somatic coliphage threshold, which was estimated for norovirus genotype I (NoV), enterovirus (EV), hepatitis A (HAV, as well as cumulatively considering the three aforementioned viruses. The dashed red line identifies the World Health Organization's annual recommended limit for the additional disease burden caused by wastewater reuse (10^{-4} DALYS per person).