Vol. 73, No. 12

Multivariate Logistic Regression for Predicting Total Culturable Virus Presence at the Intake of a Potable-Water Treatment Plant: Novel Application of the Atypical Coliform/Total Coliform Ratio[∇]

L. E. Black,[1] G. M. Brion,[1,2]* and S. J. Freitas[2]

Civil Engineering, College of Engineering, University of Kentucky, Lexington, Kentucky,[1] and College of Public Health, University of Kentucky, Lexington, Kentucky[2]

Predicting the presence of enteric viruses in surface waters is a complex modeling problem. Multiple water quality parameters that indicate the presence of human fecal material, the load of fecal material, and the amount of time fecal material has been in the environment are needed. This paper presents the results of a multiyear study of raw-water quality at the inlet of a potable-water plant that related 17 physical, chemical, and biological indices to the presence of enteric viruses as indicated by cytopathic changes in cell cultures. It was found that several simple, multivariate logistic regression models that could reliably identify observations of the presence or absence of total culturable virus could be fitted. The best models developed combined a fecal age indicator (the atypical coliform [AC]/total coliform [TC] ratio), the detectable presence of a human-associated sterol (epicoprostanol) to indicate the fecal source, and one of several fecal load indicators (the levels of *Giardia* species cysts, coliform bacteria, and coprostanol). The best fit to the data was found when the AC/TC ratio, the presence of epicoprostanol, and the density of fecal coliform bacteria were input into a simple, multivariate logistic regression equation, resulting in 84.5% and 78.6% accuracies for the identification of the presence and absence of total culturable virus, respectively. The AC/TC ratio was the most influential input variable in all of the models generated, but producing the best prediction required additional input related to the fecal source and the fecal load. The potential for replacing microbial indicators of fecal load with levels of coprostanol was proposed and evaluated by multivariate logistic regression modeling for the presence and absence of virus.

As the population increases, the burden of human waste that is put into our water systems increases, as does the potential for the transmission of enteric viruses (EV) via the water route. Of particular interest in relation to surface water systems is the control of human enteric viruses that have the ability to pass through the processes commonly employed by water treatment plants, especially under conditions of upset or peak loadings. Enteric viruses are thought to be related to endemic disease, with sporadic, undetected outbreaks of infection. It has been estimated that, annually, 10% of the U.S. population is affected by waterborne illnesses (13), and a good portion of the disease burden arises from outbreaks in which the etiological agent is not identified and may be enteric viruses. More than 100 species of viruses have been isolated from domestic sewage, and some of these have been related to waterborne outbreaks (8, 28). While most enteric virus infections may be nonsymptomatic or result in mild illness, some of the illnesses caused can be more serious and include myocarditis, infectious hepatitis, paralysis, meningitis, and severe gastroenteritis.

Providing the level of potable-water treatment required for the total elimination of all potential pathogens for the entire population is not an economically realistic goal. Common water treatment processes, with the exception of disinfection, have been designed to remove primarily sediment and other physical and chemical contaminants. Increasing the number and types of disinfection processes employed in the treatment train can improve the overall degree of viral inactivation but can cause the formation of potentially carcinogenic compounds, increase energy usage, and place a burden on budgets for facilities that are in need of other repairs. Rather than adding treatment processes and running them continuously, it is thought to be more economical to understand the behavior of enteric viruses in the environment and implement watershed management practices to prevent water sources from becoming contaminated with unacceptable levels of enteric viruses.

Presently, there are no suitable models developed that can reliably predict the presence of viable enteric viruses in surface waters. Bacterial indicators have been used to set water quality standards and regulations for body contact (33), but the relationship between bacterial indicators and the presence of enteric virus has not shown to be consistent, and waterborne outbreaks in locations where potable waters met bacteriological standards or were consistently free of indicator bacteria have been reported previously (4). Common measurements of fecal bacteria cannot specify the amounts of human and animal pollution in surface waters and therefore cannot differentiate between high and low viral risk scenarios for raw waters (15, 29). The use of alternate indicator microbes with survival characteristics similar to those of enteric viruses, such as somatic coliphages, F-specific bacteriophages, and *Bacteroides fragilis* bacteriophages, has been proposed previously (16, 19), but the relationship between the presence and absence of coliphages

---

* Corresponding author. Mailing address: 161 Raymond Building, University of Kentucky, Lexington, KY 40506-0281. Phone: (859) 257-4467. Fax: (859) 257-4404. E-mail: gbrion@engr.uky.edu.

and enteric viruses has not been consistent (29). F-specific coliphages have shown promise as a tool for identifying sources of fecal pollution (23, 24), but genotyping or serotyping with advanced techniques is required, and some researchers do not recommend coliphages as surrogate indicators for enteric viruses (9, 12).

The presence of enteric viruses is the ideal indicator of human fecal materials (4). Human enteric viruses enter surface and groundwaters through point and nonpoint sources. In water, viruses can become associated with solids suspended in the water column, and this association allows them to persist longer and deposit in sediments (4), creates problems with viral recovery from environmental samples, and influences the fate and transport of viruses in watersheds. Enteric viruses in water are detected by cell culture, PCR amplification, or a combination thereof (20, 34). These methodologies are expensive, time-consuming, and labor-intensive, and with regard to PCR, there are unresolved issues involving the potential pathogenicity of isolates (20). Other indicators of human fecal material are needed.

The fecal sterols coprostanol and epicoprostanol have been proposed as indicators of human fecal pollution. Coprostanol may constitute ~60% of the total sterols in human feces and is the primary sterol present in human sewage (25). The relationships among coprostanol, epicoprostanol, and other various sterols present in animal and human feces have been applied to the identification of fecal sources (21), with the ratio between coprostanol and epicoprostanol used to detect the presence of human fecal materials (26, 40). Unpublished research from our labs has documented epicoprostanol in a 1:100 ratio with coprostanol in raw human waste and septage.

Relying on only one or a few types of indicators for defining the fecal load or the predominant source of fecal contamination in surface waters provides an incomplete picture of the behavior of viruses in the environment. An effective model for the presence of enteric virus in surface waters requires a blend of indicators that provide information on the fecal load, the fecal source, and the amount of time aerobic processes have had to work on the fecal material, or the fecal age. The approach taken in this study uses a mixture of carefully selected indicators for fecal load, source, and age for modeling the presence of viable enteric virus in raw waters. This new paradigm for modeling the presence of pathogens provides a basis on which to build future watershed protection and management models.

The ratio of atypical coliforms (AC) to total coliforms (TC), a number calculated by dividing the concentrations of bacteria that produce atypical colonies on membrane filters supplied with total coliform broth by the concentrations of bacteria that produce typical TC colonies or AC colonies, is a crucial input parameter that accounts for fecal age and modifies information provided by load and source indicators. The term parameter refers to explanatory environmental variables which would relate to the dependent variable, here the presence of enteric virus, measured for incorporation into models. This is the first study designed to collect information on the AC/TC ratio relative to the presence of enteric virus. Freitas et al. (14) completed a preliminary study that showed the AC/TC ratio to be the most essential parameter for predicting the presence of viable enteric virus and supported the research design. The present study proposes that a select combination of indicators that represent the fecal load (levels of fecal coliforms, *Escherichia coli* bacteria, and coprostanol), fecal source (levels of male-specific coliphages and epicoprostanol), and fecal age (the AC/TC ratio), along with parameters (e.g., temperature) that are related to the stability of enteric virus in environmental waters, will be effective in modeling the presence of enteric virus in the Kentucky River. The findings relating enteric virus presence to the novel AC/TC ratio presented herein are derived from a database compiled in response to a completed EPA-funded study (R-82978401). The database is available through NCER at http://es.epa.gov/ncer/index.html, which is searchable by grant number.

The Kentucky River is a tributary of the Ohio River, 259 mi (417 km) long, whose watershed encompasses 7,000 mi^2 (18,000 km^2) and which flows south to north through central Kentucky in the United States (http://kywater.org/watch/ky .htm). The watershed lies above thick layers of easily dissolved limestone that form carbonate aquifers. The Kentucky River is used for recreational purposes (e.g., swimming and fishing) and is a source of potable water, providing more than 95% of the drinking-water supply to the 600,000 people living along its basin (39). The river is divided along its entire length into 14 locks, which create a series of pools that act as individual shallow lakes, allowing the resuspension of deposited sediments during periods of high flow. Central Kentucky receives >100 cm of precipitation annually, often in intense storms. Samples from the Kentucky River were collected at the inlet of a potable-water treatment plant (Frankfort Plant and Water Board) located between lock no. 5 and lock no. 4. A major tributary into this pool comes from the Glenn's Creek watershed (16 mi^2, or 21,615 acres), which empties very near the intake. Glenn's Creek watershed land usage is classified as 85% agricultural, 5% wooded and rural, and about 10% residential, commercial, or industrial (http://kywater.org/watch/ky .htm). Three businesses or organizations hold permits to release discharges into the major creeks of the Glenn's Creek watershed, Camden Creek and Buck Run. One of these businesses is a sewage treatment plant for the city of Versailles. Fecal contamination from both animal and human sources in the river pool at a point prior to the inlet has been documented previously (http://kywater.org/watch/ky.htm), and sewage discharges from the cities of Wilmore, Lexington, Lawrenceburg, and Nicholasville enter the river from tributaries above the inlet. The raw-water quality at the inlet is quite volatile and fluctuates with storm events along the entire length of the river, making water quality modeling a complex problem.

MATERIALS AND METHODS

Surface water samples and processing. Surface water samples from the Kentucky River were collected at the inlet of a drinking-water treatment facility, Frankfort Plant and Water Board, from June 2003 to March 2005. Physical, chemical, and biological raw-water quality parameters were analyzed and recorded for each sample. The physical parameters considered were turbidity, temperature, the amount of rainfall, and flow. The alkalinity, pH, and hardness, along with the levels of the fecal sterols coprostanol and epicoprostanol, were the chemical parameters measured. The biological parameters measured were the levels of AC, fecal coliforms, TC, *E. coli* bacteria, enterococci, *Giardia* species cysts, *Cryptosporidium* species oocysts, EV, somatic coliphages, male-specific coliphages, *Clostridium perfringens* spores, and male-specific salmonella phages. The chemical and biological parameters were analyzed at the Environmental

Research and Training Laboratories at the University of Kentucky, Lexington, by previously published methods.

Enteric virus collection and detection. Enteric viruses were recovered from river water by the virus adsorption-elution (VIRADEL) methodology and assayed on buffalo green monkey kidney (BGMK) cells to determine the most probable number (MPN) per 100 liters as specified by the U.S. Environmental Protection Agency (EPA) (34). Briefly, at least 200 liters of river water was passed through positively charged VIRADEL 1MDS filters at the sampling site. An eluent comprising beef extract (GIBCO) at pH 9.5 was used to elute the filter. The eluate pH was reduced to 3.5, and the eluate was allowed to flocculate for 3 h. The floc was pelleted by centrifugation at $4,000 \times g$ for 20 min. The supernatant was discarded, and the pellet was resuspended in a solution of sodium hydrogen phosphate (pH 9.5). The resuspended pellet solution was stirred gently for 10 min and then recentrifuged at $4,000 \times g$ for 20 min, after which the pellet was discarded while the supernatant solution was retained. The pH of the supernatant was adjusted to 7.0 to 7.4 by using dilute hydrochloric acid and sodium hydroxide, after which the supernatant was subjected to sterile filtration through a series of beef extract-pretreated 0.8-, 0.4-, and 0.22-μm-pore-size sterile syringe filters. The final concentrate volume (~30 ml) of the sample was noted, and the inoculum volume of the sample was calculated using the EPA formula. The inoculum sample was apportioned for cell culture, with the remainder archived at −80°C. BGMK cells grown on the flat side of 25-cm^2 sterile plastic flasks for 3 to 5 days until a cell monolayer formed were used for the detection of the presence of enteric virus by cytopathological effects (10). Ten flasks were inoculated with the precalculated volume of the sample inoculum. The flasks were laid flat, with intermittent rocking, for 120 to 150 min before overlay medium (50% minimal essential medium with Hanks salts and L-glutamine [GIBCO], 50% Leibovitz L-15 medium with L-glutamine [GIBCO], 10% fetal bovine serum, and 10.8 mM NaHCO$_3$) was added. These flasks were then placed into an incubator set at 36.5°C. For the next 14 days, the flasks were checked for cytopathic effects and/or the disappearance of the cell monolayer. Cytopathic effects indicated the presence of enteric virus. Readings for all flasks were confirmed by a second passage and, when required by conflicting results, a third passage. MPN estimates of viral concentration were calculated, and the detection limit was 2.1/100 liters.

Bacterial, protozoan, and coliphage analyses. Total, atypical, and fecal coliforms were detected by membrane filtration of appropriate dilutions of samples per standard methods (1). *E. coli* bacteria and enterococci in 100-ml samples were detected by Colilert and Enterolert Quanti-Tray/2000 (IDEXX Laboratories Inc., Westbrook, ME) methods, respectively. *Cryptosporidium* oocysts and *Giardia* cysts in 10-liter samples were detected using EPA method 1623 for filtration, immunomagnetic separation, and fluorescent-antibody microscopy (38). Somatic coliphages (φX-174 was used as a control) and male-specific coliphages (MS2 was used as a control) were detected by the EPA single-agar-layer coliphage method 1602 (37), with at least 200 ml of water assayed for male-specific coliphages.

Fecal sterols. Black and Brion (L. E. Black and G. M. Brion, presented at the 13th IWA Conference on Health-Related Water Microbiology, Swansea, United Kingdom, September 2005) described a quantitative methodology for analyzing fecal sterols by gas chromatography-mass spectrometry (GC-MS) after concentration and extraction from 3-liter samples by solid-phase techniques (2). A J. T. Baker BAKERBOND Speedisk extraction manifold equipped with a BAKERBOND Speedisk C$_{18}$ disk (50 mm in diameter; Mallinckrodt Baker, Inc., Phillipsburg, NJ) was assembled. The Speedisks were sequentially preconditioned with 30 ml of dichloromethane (CH$_2$Cl$_2$; Chemical Abstract Service [CAS] no. 75-09-2; Sigma-Aldrich, St. Louis, MO), 30 ml of acetone (CH$_3$COCH$_3$; CAS no. 67-64-1; Sigma-Aldrich, St. Louis, MO), two 30-ml aliquots of acetonitrile (CH$_3$CN; CAS no. 75-05-8; Sigma-Aldrich, St. Louis, MO), and two 30-ml aliquots of water (Optima grade; Sigma-Aldrich, St. Louis, MO) without drying. The 3-liter environmental samples were then vacuum filtered, followed by a wash with 30 ml of H$_2$O-CH$_3$CN (80:20, vol/vol) solution. The laden sample disks were allowed to dry under a vacuum, and then sterols were eluted with 10 ml of CH$_3$COCH$_3$ and 20 ml of CH$_2$Cl$_2$. A glass funnel was plugged with glass wool (Ohio Valley Specialty Chemical, Marietta, OH) and filled with anhydrous sodium sulfate (Mallinckrodt Baker, Inc., Phillipsburg, NJ). The eluate was funneled through the sodium sulfate to remove any residual water into an evaporation container (Labconco Rapid Vap, Kansas City, MO) and condensed under nitrogen at 35°C to ~1 ml. The condensed eluate was transferred to an 8-ml glass analysis vial (Kimble, Vineland, NJ) and dried under a gentle nitrogen stream. The eluate residue was reconstituted with pyridine-BSTFA [bis(trimethylsilyl)trifluoroacetamide; 1:1, vol/vol] solution for derivatization, and an internal standard, chrysene-d$_{12}$, was added for quantification by GC-MS analysis. Analysis by GC-MS was completed

on a Varian CP-3800 gas chromatograph equipped with a Varian Saturn 2200 quadrupole ion trap mass spectrometer.

Statistical analysis. A database of parameter readings from 108 sampling events was assembled. Of these, only the 100 readings that included information on all 18 parameters were used for modeling and descriptive statistics. Data recorded as being above an upper detection limit were assigned the value of the detection limit for the purposes of calculating averages, producing transformations of data, and modeling. Data below the detection limits were assigned a value that would provide separation from measurements close to the detection limits. Bacterial parameters (e.g., levels of AC, TC, fecal coliforms, *E. coli* bacteria, and enterococci) and the bacteriophage parameter (the level of somatic coliphages) were logarithmically transformed to calculate geometric means and to create a more normal distribution for modeling as per standard convention, with the few readings that were below the detection limit assigned an initial value of 0.05 organisms per 100 ml before transformation. Measurements of epicoprostanol, *Cryptosporidium* oocyst, EV, and male-specific coliphage levels were dichotomized as 0 for nondetection when measurements were below detection limits and 1 when measured levels were above detection limits due to the large number of nondetection observations, the frequency of measurements at or near the lower detection limits when detectable levels were present, and an understanding of the expected rates of recovery by the methods employed. Measurements of coprostanol and *Giardia* cyst levels were first assigned values of 0.33 pg/ml and 1 cyst/10 liters, respectively, if measurements were below the detection limits for these parameters; then these assigned values were divided by the maximum value observed for the respective parameter to obtain signals that optimized model fitting based on past experience. The remaining parameters (turbidity, temperature, pH, alkalinity, and hardness) were not transformed as either the variances in measurements were not large or the values were found to be normally distributed.

SigmaStat version 2.03 was used to perform the multivariate logistic regressions, Mann-Whitney rank sum tests, and Pearson correlations. For comparisons between two dichotomous variables (e.g., the levels of EV and epicoprostanol), the chi-squared coefficient corrected for continuity as per Yates was used to evaluate relatedness and the level of significance. For the first model, parameters that showed significant correlation with the presence of enteric virus at a level of significance of *P* of <0.05 (turbidity, temperature, alkalinity, coprostanol and epicoprostanol concentrations, the AC/TC ratio, and levels of *E. coli* and *Giardia* organisms) were used to fit a multivariate logistic regression model. Then the selection and substitution of input parameters according to the proposed age, load, and source paradigm were performed, and parameters were evaluated in subsequent models.

RESULTS

Descriptive statistics. The descriptive statistics for the physical, chemical, and biological parameters measured from June 2003 to March 2005 for the Kentucky River are listed in Table 1. For variables that were not logarithmically transformed for modeling, an arithmetic mean and a corresponding standard deviation were calculated and recorded, along with the maximum and minimum values. For variables that were logarithmically transformed for modeling, a geometric mean and its corresponding geometric standard deviation were calculated and recorded, along with the maximum and minimum values. For the dichotomized variables, the minimum and maximum values were recorded along with the number of times that the indicators were detected in the river samples.

The average alkalinity and hardness values and pHs reflect the degree of limestone saturation in the river system from the surrounding karst under normal conditions. The pH of the river was unaffected, even when substantial amounts of runoff generated by rainfall dropped the alkalinity to low levels (16 mg/liter as CaCO$_3$), due to the buffering capacity present. However, the wide range in turbidity (4.2 to 662.0 nephelometric turbidity units) shows that raw-water quality was significantly affected by rainfall events, with the river picking up increased particle loads from runoff, a condition that was

TABLE 1. Descriptive statistics of the physical, biological, and chemical parameters for the Kentucky River from June 2003 to March 2005

Parameter type	Parameter	Arithmetic mean (SD), geometric mean (GSD),[b] or frequency of detection	Minimum	Maximum
Physical	Turbidity (NTU[a])	61.9 (98.14)	4.2	662.0
	Temp (°C)	17.4 (7.07)	4.3	27.2
Chemical	pH	7.8 (0.17)	7.2	8.4
	Alkalinity (mg/liter as $CaCO_3$)	93.3 (21.04)	16.0	130.0
	Hardness (mg/liter as $CaCO_3$)	138.4 (30.43)	64.0	240.0
	Coprostanol content (ng/liter)	18.4 (22.25)	<3.3[c]	110.8
	Epicoprostanol content (ng/liter)	Detected in 32 of 100 samples	<3.3[c]	49.5
Bacterial	Level of:			
	Atypical coliforms (CFU/100 ml)	**6,754 (4.7)**	128.0	355,000
	Total coliforms (CFU/100 ml)	**458 (4.8)**	<1[d]	53,182
	AC/TC ratio	20.7 (16.99)	1.5	74.3
	Level of:			
	Fecal coliforms (CFU/100 ml)	**61 (7.4)**	<1[d]	2,033
	E. coli bacteria (MPN/100 ml)	**76 (5.0)**	2	>2,419[e]
	Enterococci (MPN/100 ml)	**53 (6.8)**	1	>2,419[e]
Protozoan	No. (per 10 liters) of:			
	Giardia cysts	5 (5.9)	<1[d]	35
	Cryptosporidium oocysts	Detected in 46 of 100 samples	<1[d]	16
Viral	Level of:			
	EV (MPN/100 liters)	Detected in 58 of 100 samples	<2.1	71
	Somatic coliphages (PFU/100 ml)	**40 (5.2)**	<1[d]	1,616
	Male-specific coliphages (PFU/100 ml)	Detected in 74 of 100 samples	<1[d]	275

[a] NTU, nephelometric turbidity units.
[b] Geometric mean values are in bold. GSD, geometric standard deviation.
[c] The detection limit for fecal sterols was <3.3 ng/liter.
[d] The detection limit is accepted to be 1 organism per standard volume.
[e] The IDEXX method had an upper detection limit of 2,419/100 ml.

sometimes exacerbated by sediment resuspension from the series of pools created by the river's lock and dam system. On the days of greatest turbidity, the river looked like flowing milk chocolate. The river's temperature mimics the seasonality of the region, with extreme temperature changes between the warm summer months (27.2°C) and cold winter months (4.3°C). Temperature, of course, impacts microbial growth and survival, with colder temperatures inhibiting growth and prolonging pathogen survival.

EV was detected in only 58 of the 100 samples and was present in low densities. Only 19 of the 58 virus-positive samples had MPNs of ≥10/100 liter. The maximum MPN recorded for EV was 71/100 liter, on 1 July 2003, when the recorded temperature of the river was 22.1°C. Sixteen of the 19 days on which EV levels were above an MPN of 10/100 liters occurred between May and October, times when the river was used for recreational purposes and the temperature averaged 21.6°C and ranged from 16.0 to 27.2°C.

Microbial indicators of fecal contamination were present, with concentrations of enterococci consistently higher than the recommended recreational-contact standard of 33 CFU/100 ml, as reflected in the geometric mean of 53 CFU/100 ml. Of the samples, 26% had concentrations higher than the single maximum allowable enterococcus density for infrequent full-body-contact recreation as recommended by the EPA (151 CFU/100 ml). The *E. coli* concentrations in 18% of the samples were greater than the recommended 576-CFU/100-ml single maximum allowable *E. coli* density for infrequent full-body-

contact recreation as recommended by the EPA, even though the geometric mean was lower than the recommended 126 CFU/100 ml, at 76 CFU/100 ml. During storm events, the levels of *E. coli* and enterococci rose to >2,400 CFU/100 ml three and seven times, respectively.

The microbial data from the river can be used to indicate the relative amounts of agricultural and human sewage entering the river and provide an indication of the age of these fecal materials. The mean AC/TC ratio (20.7) was indicative of contamination from aged agricultural and suburban runoff when compared to runoff values from other studies in the area (3, 7, 31). The high AC/TC ratio showed the predominance of indigenous, nutrient-linked bacteria over those introduced from fecal sources. The lowest AC/TC ratio (1.5) was suggestive of the presence of fresh human sewage, which may be input from combined storm effluents and bypassed flows released from upstream wastewater treatment plants. Male-specific coliphages were present 74% of the time and generally in low densities; 44 of the 74 readings detected >2 PFU/100 ml even though 200 ml of the sample was assayed. The indicators more specific for human waste, epicoprostanol and EV, were present only approximately 50% of the time, suggesting that the river was predominantly impacted by older agricultural runoff mixed with human input. Based on this information and the observed *E. coli* and enterococcus levels, recreational use of the river involving full-body contact should be avoided, especially in the summer months.

TABLE 2. Chi-square and Pearson values of correlation, along with levels of statistical significance, between the presence of EV and other measured parameters[a]

Parameter	Value of correlation with the presence of EV	P value
Turbidity	0.235	0.018
Temp	−0.551	<0.001
pH	0.038	0.709
Alkalinity	−0.218	0.029
Hardness	−0.161	0.109
Coprostanol content*	0.258	0.010
Epicoprostanol content*	**9.086**	**0.003**
Level of:		
AC*	−0.140	0.166
TC*	0.142	0.158
AC/TC ratio*	−0.513	<0.001
Level of:		
Fecal coliforms*	0.186	0.064
E. coli bacteria*	0.255	0.011
Enterococci*	0.116	0.252
Giardia cysts*	0.315	0.001
Cryptosporidium oocysts*	**1.314**	**0.252**
Somatic coliphages*	0.130	0.196
Male-specific coliphages*	**0.037**	**0.846**

[a] Asterisks indicate transformed variables. Chi-square values are in bold.

Parameter correlation. The Pearson correlation coefficients and chi-squared values with their respective levels of significance for correlations between EV and the other measured parameters are listed in Table 2. The parameters of turbidity, temperature, alkalinity, coprostanol and epicoprostanol concentration, AC/TC ratio, and E. coli and Giardia level were significantly correlated with the presence of EV. Among the significantly correlated variables, the fecal age-related variables of AC/TC ratio and temperature were closely correlated ($P < 0.001$) with the presence of EV and with each other. Larger amounts of EV, due to the presence of raw human sewage, correlated with a low AC/TC ratio (<1), resulting in the expected negative correlation. An indirect indicator of fecal age, turbidity, was also strongly negatively correlated with the presence of EV but was not correlated with the AC/TC ratio. Least strongly correlated with the presence of EV were the level of male-specific coliphages and the pH (P, 0.846 and 0.709, respectively). Although the presence of male-specific coliphages has been suggested by others as a good source indicator for human fecal materials and enteric viruses, an alternative fecal source indicator, the presence of epicoprostanol, had a more statistically significant chi-squared relationship ($P = 0.003$) with the presence of EV than the presence of male-specific coliphages in this watershed. Other parameters that were more directly related to fecal materials and could be used as indicators of the fecal load (levels of Giardia cysts, E. coli bacteria, and coprostanol) were significantly correlated with the presence of EV. These initial correlations provided a starting point for selecting input parameters for modeling.

Multivariate logistic regression. Initially, to assist in selecting input parameters for inclusion in and exclusion from models for the prediction of the presence of EV, all parameters with correlation coefficients or chi-squared statistics with P of <0.05 were use as input variables in the model (model 1) (Table 3). This selection method reduced the potential input variables from 17 to 8, and the results are tabulated in Table 3. Although the correlation coefficients were helpful, in order to optimize the modeling approach, smaller models based upon input parameters selected from the initial eight were evaluated in an attempt to determine an optimum combination of three input parameters reflecting the load, source, and age that could be easily measured by the local water utility. Not all of the models that were created and evaluated are presented in Table 3, only those that provided good fits to the data as indicated by a prediction of virus presence equivalent to or better than that given by model 1.

Of the highly correlated input parameters, model 1 found turbidity, alkalinity, coprostanol and epicoprostanol concentrations, and the level of E. coli bacteria to be minimally significant, as the small coefficients and high P values indicate. The AC/TC ratio, temperature, and the concentration of Giardia cysts were the most significant ($P < 0.1$) input variables in model 1. Subsequent models 2 to 5 were created by a process that did not rely solely on the calculated value of parameter significance but used a paradigm of selecting input parameters to indicate conditions relative to the fecal age (the AC/TC ratio), fecal source (the epicoprostanol concentration), and fecal load (the levels of Giardia cysts, E. coli bacteria, fecal coliforms, and coprostanol) based on expert judgment and an analysis of data quality.

Four three-input-parameter models were created, and their performances in predicting the positive and negative EV readings on which the models were fit were evaluated. These models were created by fixing age and source indicators as the AC/TC ratio and the presence of epicoprostanol and then varying the load input parameter. The model that predicted the most positive readings was model 4, in which the AC/TC ratio was combined with the presence or absence of epicoprostanol and the density of fecal coliforms. A total of 49 (84.5%) of 58 positive readings were correctly identified by model 4, with 33 (78.6%) of 42 negative readings correctly identified. Models 2, 3, and 5 used different fecal load indicators (the concentrations of Giardia cysts, E. coli bacteria, and coprostanol, respectively), parameters that were more significantly related to the presence of EV than the density of fecal coliforms according to a simple correlation analysis (Table 2). However, these models did not perform as well as model 4, which used the load signal of fecal coliforms. The enhanced performance of model 4 in comparison to those of the other models is thought to be due to the quality and completeness of the fecal coliform data, as every sampling event produced a recordable observation within both upper and lower detection limits. Coprostanol data were truncated at the lower end of the detection range (<3.3 ng/liter), and coprostanol was not detected at all in 29 samples. Measurements of E. coli concentrations were truncated at the upper end of the range. All three-parameter models performed similarly in their ability to predict the absence of EV, with the exception of model 5. In model 5, the coefficient value fit to the dichotomous variable of the epicoprostanol level was lower than the coefficient values assigned to variables in models 1 to 4, and this may account for

TABLE 3. Multivariate logistic regression models for the prediction of the presence of enteric viruses

Model	Input variables	P value[a]	% (No.) of correct positive predictions[b]	% (No.) of correct negative predictions[c]	Multivariate logistic regression equation[d]
1	Turbidity	0.407	**79.3** (46)	**76.2** (32)	$\text{Logit}(P) = 7.401 + (0.00481 \cdot \text{TURB})$ $- (0.140 \cdot \text{TEMP}) - (0.0137 \cdot \text{ALK})$ $+ (2.080 \cdot \text{COP}) - (0.530 \cdot \text{EPI}) -$ $(3.190 \cdot \text{AC/TC}) - (0.0796 \cdot \text{ECOLI})$ $+ (4.806 \cdot \text{GIA})$
	Temp	**0.006**			
	Alkalinity	0.424			
	Coprostanol content	0.327			
	Epicoprostanol content	0.567			
	AC/TC ratio	**0.003**			
	Level of *E. coli* bacteria	0.896			
	Level of *Giardia* cysts	0.098			
2	AC/TC ratio	**<0.001**	**82.8** (48)	**76.2** (32)	$\text{Logit}(P) = 3.787 - (3.382 \cdot \text{AC/TC}) +$ $(0.685 \cdot \text{EPI}) + (4.371 \cdot \text{GIA})$
	Epicoprostanol content	0.261			
	Level of *Giardia* cysts	0.083			
3	AC/TC ratio	**<0.001**	**81.0** (47)	**78.6** (33)	$\text{Logit}(P) = 2.840 - (3.355 \cdot \text{AC/TC}) +$ $(1.099 \cdot \text{EPI}) + (0.688 \cdot \text{ECOLI})$
	Epicoprostanol content	0.062			
	Level of *E. coli* bacteria	0.078			
4	AC/TC ratio	**<0.001**	**84.5** (49)	**78.6** (33)	$\text{Logit}(P) = 3.030 - (3.331 \cdot \text{AC/TC}) +$ $(1.260 \cdot \text{EPI}) + (0.552 \cdot \text{FC})$
	Epicoprostanol content	**0.036**			
	Level of fecal coliforms	0.136			
5	AC/TC ratio	**<0.001**	**84.5** (49)	**69.1** (29)	$\text{Logit}(P) = 4.414 - (3.816 \cdot \text{AC/TC}) +$ $(0.223 \cdot \text{EPI}) + (3.393 \cdot \text{COP})$
	Epicoprostanol content	0.764			
	Coprostanol content	0.113			

[a] Values in bold are statistically significant at a 0.05 probability level.

[b] The percentages of correct positive predictions were calculated by dividing the numbers of correct positive predictions by the total number of actual positive readings (58).

[c] The percentages of correct negative predictions were calculated by dividing the numbers of correct negative predictions by the total number of actual negative readings (42).

[d] $y = 1/\{1 + \exp[-\text{logit}(P)]\}$. TURB, turbidity; TEMP, temperature; ALK, alkalinity; COP, coprostanol concentration; EPI, epicoprostanol concentration; AC/TC, AC/TC ratio; ECOLI, level of *E. coli* bacteria; GIA, level of *Giardia* cysts.

the model's lesser ability to distinguish samples in which EV was absent.

DISCUSSION

The combination of fecal load, source, and age parameters used in model 4 effectively captured the relationship between enteric virus presence and water quality in this river system, with the AC/TC ratio driving model performance. Model 4, which used the AC/TC ratio, the level of epicoprostanol, and the fecal coliform concentration as input variables, obtained the most correct identifications of positive (85.4%) and negative (78.6%) readings. The model was simple and is an approach that could be implemented by the local water utility. Although models 1 and 2 had performances similar to that of model 4, with only slightly fewer positive readings predicted accurately, these models would not be readily implemented by the water utility. Model 1 is overly complex, and model 2 relied upon measuring concentrations of a difficult-to-assay pathogenic protozoan to indicate the fecal load. The replacement of the level of *E. coli* bacteria with the fecal coliform concentration as an input parameter to indicate the fecal load worked because these two indicators are closely correlated ($P < 0.001$), as would be expected. The correlation between the presence of fecal coliforms and that of EV ($P = 0.064$) was nearly at the P level of <0.05 used to initially select model input parameters. What is clear from all of the models created is that the underlying paradigm of selecting indicators associated with the fecal age, load, and source is appropriate and that the AC/TC ratio

is the key to success for this watershed. The multiparameter paradigm developed in this study avoided the problems associated with "silver-bullet" approaches that rely on single indicators of only the source or the load and added information on fecal age related to pathogen survival.

The importance of including an indicator of fecal age (the AC/TC ratio) for modeling virus presence was highlighted by the results presented above. This finding is not unexpected. In other studies, the AC/TC ratio has been able to distinguish between raw human sewage (0 to 5), flowing agricultural runoff (10), flowing urban runoff (20), and aged urban runoff (>100) with statistical significance, and the ratio consistently rises as fecal material is aged under aerobic conditions in the environment (6, 7, 31). It has been suggested previously that the AC/TC ratio could serve as a watershed quality standard for the Kentucky River, with values greater than 20 associated with high fecal coliform levels and enhanced pathogen risk (31). Simply sorting the database used for modeling by ascending AC/TC ratios highlights the utility of this novel indicator. If the AC/TC ratio was measured at less than 15, which is lower than the average value of 20.7 for the river and therefore indicative of the presence of fresher-than-usual fecal material, the frequency of EV detection increased. Of the 51 samples in which the AC/TC ratio was less than 15, 74% were positive for enteric viruses. This statistic represents 43 of the total of 58 EV-positive readings recorded. The relationship between the presence of enteric viruses and the AC/TC ratio is easily seen in Fig. 1, in which increasing ratios are associated with the ab-
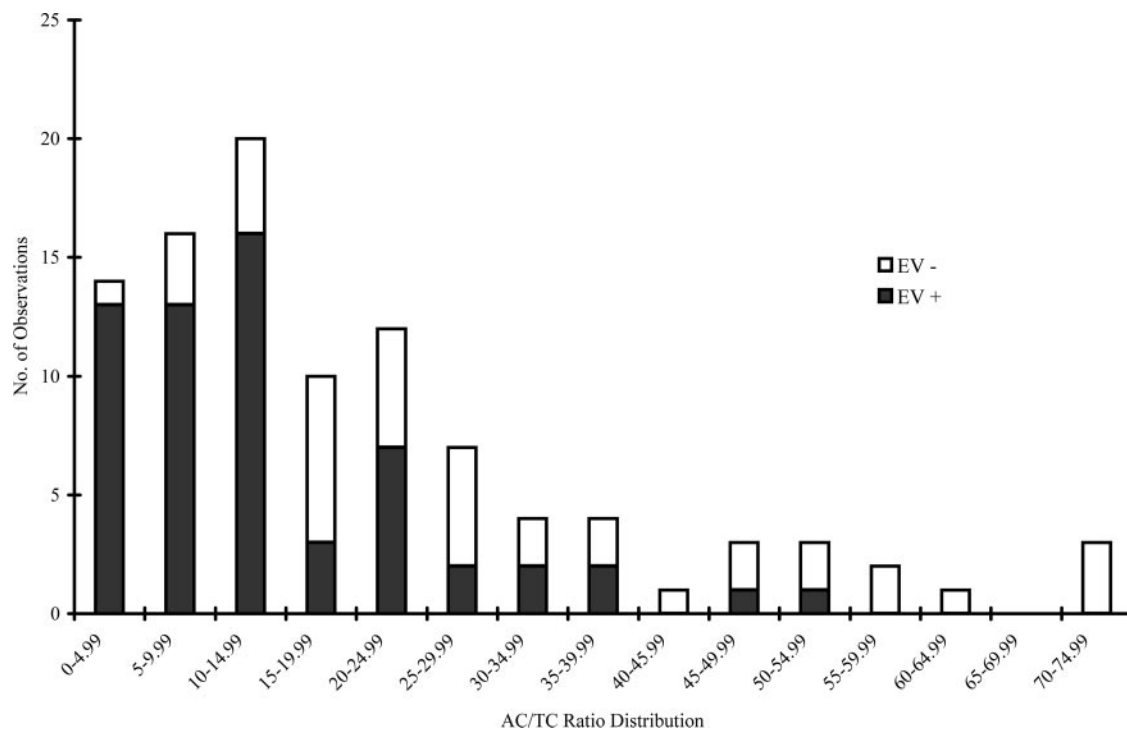
FIG. 1. AC/TC ratio distribution as related to the presence or absence of EV.

sence of EV. There were 14 times when the AC/TC ratio was below 5, a condition indicative of agricultural wastes being diluted with generous amounts of human sewage, and in all but one of these cases, enteric viruses were detected in the raw water. What is very clear is that decreased AC/TC ratios, those below the average of 20, were associated with the increased presence of EV. The predictive ability of all models relied upon the ability of the AC/TC ratio to represent fecal age as well as to provide some insights about the fecal source. Further logistic modeling efforts (data not shown) using only the AC/TC ratio to predict the presence of EV had a 77.6% success rate. This represents only one less positive reading than the number predicted by model 1. The prediction of the absence of EV suffered (64.3%) when only the AC/TC ratio was used. Clearly, other input parameters in combination with the AC/TC ratio are needed to reliably predict the presence and absence of enteric viruses in water, input parameters that can give more information on the fecal load and source.

These models may be improved by the addition of indices that can provide information on degrees of change and the potential for pathogen survival in the watershed. Temperature was significantly correlated with the presence of EV ($P <$ 0.001), as expected. Low temperatures during the winter months may enhance viral survival and suppress aerobic microbial degradation. Higher temperatures were associated with summer days when the MPN of EV present was >10/100 liters. These seasonal effects are incorporated into the AC/TC ratio, which is a measure of the ratio of indigenous bacterial activity to that of fecally associated bacteria introduced into the watershed. This incorporation of seasonal variations was reflected in the significant correlation between the temperature and the AC/TC ratio ($P <$ 0.001). Removing temperature as an input

parameter from the multivariate logistic regression model did not negatively impact the predictions; actually, the simpler models performed slightly better.

Commonly measured physical and chemical parameters may not always be associated with increased viral presence. Turbidity is related to the fecal load from storm input, as well as fecal age, and these relationships are reflected in the values of correlation with the presence of EV. However, changes in turbidity can also be related to algal growth in the Kentucky River or sediment input not related to human fecal materials. Therefore, the inclusion of this variable may confound the signals from other input parameters. Excluding turbidity from models actually improved predictions. This finding is supported by the results of other studies that have shown turbidity to be of little utility for predicting the presence of protozoan pathogens (30). More research needs to be done to create new and better indicators of change and fecal age, expanding upon the demonstrated utility of the AC/TC ratio.

Although bacterial indicators have traditionally been used to assess human health risk in waters, they do not adequately distinguish between animal and human contamination and are not in themselves indicative of the risk inherent from enteric viruses in our surface waters. The presence and density of bacteria can provide some indication of the presence of viruses when the contamination originates from human sources; however, this relationship may not exist when the source of pollution is of animal origin (11, 32). Viruses have been found in the environment in the absence of traditional bacterial indicators, especially coliform bacteria, but virus collection and assay methods are out of the reach of ordinary labs, and some type of easily measured indicator is needed, even if it is an imperfect one. There has been a great deal of work done to identify other

indicators for enteric viruses from among other fecally associated microbes and chemicals. Enterococci, *C. perfringens* spores, and bacteriophages (somatic coliphages, male-specific F RNA bacteriophages, and *Bacteroides* phages) have been proposed previously (17, 18), but the validity of such indicators has not been conclusively proven. Standardization and quality control in the United States have produced standard methods for coliphages (37) and enrichment methods that provide for the better detection of small quantities of coliphages (36), but our results do not support the use of male-specific coliphages for modeling the presence of enteric viruses for this river system.

Male-specific coliphages, although detected more frequently than EV, were not statistically related to the presence of EV in the raw waters by chi-square analysis ($P = 0.846$). Samples from the river were generous (200 ml or more), but the numbers of male-specific coliphages recovered by the single-layer assay were low. Other research on urban and agricultural runoff in a watershed that feeds the Kentucky River upstream of the intake found only 25% of 150 2-liter animal-impacted runoff samples concentrated with polyethylene glycol precipitation to have detectable levels of male-specific coliphages (5). The prevalence of male-specific coliphages in the Kentucky River is greater than that reported in this prior runoff study by a factor of three, suggesting input from sources such as domestic sewage, but male-specific coliphage densities were not strongly correlated with those of the indicator bacteria, suggesting that the bulk of fecal material in the river is from animal sources mixed with some human sewage.

Others have found that levels of coprostanol are directly related to primary and secondary contact limits for the density of fecal coliforms (22). Our study supports this prior finding, with increasing levels of coprostanol associated with above-average levels of *E. coli* contamination. The geometric mean concentration of *E. coli* bacteria in samples in which the levels of coprostanol were >18 ng/liter was an MPN of 144/100 ml, well over the EPA-recommended standard of 126 CFU/100 ml. The geometric mean concentration in samples in which the observed coprostanol level was <18 ng/liter was an MPN of only 56/100 ml. The two geometric means were found to be significantly different from each other by the Mann-Whitney rank sum test ($P = 0.004$). It would seem that coprostanol is a potentially useful indicator for defining fecal load categories. If coprostanol levels were used in conjunction with epicoprostanol levels in a ratio, or with additional indicators of human fecal sources, a good representation of the fecal load and the predominant source could be obtained. Analytical methods recently developed for the capture and detection of environmental sterols use an automated extraction process that greatly simplifies this analysis, increases recovery, and speeds up sample processing time to 4 h, allowing for faster responses to watershed quality changes.

The presence of detectable levels of epicoprostanol was more closely related to the presence of EV than that of male-specific coliphages in this watershed. This result is to be expected in light of the small amounts of epicoprostanol that would be expected to come from animal sources of fecal materials. Male-specific coliphages in amounts detectable in environmental water samples have been associated with animal sources (23), and this finding may confuse the issue of utilizing

them as indicators of strictly human sources without further classification into genotypes. Levels of epicoprostanol in animal fecal materials have been documented to be much lower than those of other sterols, such as coprostanol (21), so the presence of epicoprostanol in environmental samples is thought to be indicative of human fecal materials. The strength of correlation between epicoprostanol levels and the presence of enteric virus, and the ability to use coprostanol as a load signal, should spur continued interest in the development of methods for using fecal sterols as indicators of the fecal source and load relative to the presence of enteric virus.

Improvements in extraction methods for fecal sterols may increase their utility as indicators of human fecal material and enteric viruses in surface waters, especially with respect to epicoprostanol. In our study, the levels of epicoprostanol were low and often undetectable (<3.3 ng/liter with extraction and detection combined), as would be expected from a watershed impacted primarily by agricultural input. Only 32 of 100 samples gave measurable levels. The method used for this study was based on that of other studies that used solid-phase extraction (2) and provided a rate of recovery of sterols from duplicate sterol-spiked river samples of <30% (unpublished data). The application of a simplified, commercially available Soxhlet extraction method for sterols has shown rates of recovery of epicoprostanol from spiked environmental samples nearly three times higher than those shown by the solid-phase extraction method used in this study (unpublished data). It is conceivable that using a different extraction method would have allowed for the calculation of a direct ratio of coprostanol to epicoprostanol, a relationship found to be applicable to the apportionment of fecal sources, and provided better results, particularly for the prediction of the absence of EV.

This modeling exercise points out issues that are encountered with data quality with respect to the method used to detect enteric viruses. All five of the models developed predicted the same six samples to be positive for viral presence when the database had recorded viral absence. These mispredictions may actually be indicative of samples in which the enteric virus MPN method missed detecting the presence of virus due to a low level of recovery by the standardized method. Viral presence and absence were confirmed for all samples, but losses at the concentration steps may have prevented detection. Of these six readings, five of them were made on days when the AC/TC ratio was <15, indicative of the presence of fresher-than-normal fecal material in the river and a condition strongly associated with viral presence. Epicoprostanol was detected in four of the samples and male-specific coliphages were detected in five of the six samples, conditions that would seem to indicate the meaningful presence of a large load of human wastes in the system. For three of the six readings, the levels of *E. coli* bacteria were above the geometric mean MPN of 76/100 ml, indicating a larger than normal fecal load. This combination of measurements was not normally associated with the absence of enteric viruses. It may be that these readings recorded false-negative results for the presence of enteric virus. This scenario is plausible considering the limitation of the ability of the VIRADEL-organic flocculation method to recover viruses from large volumes of complex environmental matrices and the fact that in the MPN method, not all of the sample is analyzed. In prior testing of the VIRADEL-

organic flocculation method for the collection and reconcentration of large volumes of water spiked with enteric viruses, relatively low and quite variable levels of recovery efficiency were documented (27). It is also known that when filtering to remove toxicity before cell culture, a 30% loss of signal can be expected (35). Solids in raw waters can interfere with recovery rates as well. Our own internal quality assurance and quality control procedures documented an average percentage of recovery of 35% for matrix-spiked samples, with initial precision and recovery rates for five separate laboratory control sample preparations of 38% for the plaque assay and 28% for the MPN method. With the wide variety of raw-water conditions we faced and the known and expected losses of viral signal due to the method used, it is plausible that virus presence was not always detected and that the readings in question may well have been recorded as false negatives. Although all five models consistently missed predicting the presence of enteric virus in the same three samples, a review of the data does not provide any additional insight into these mispredictions, and these observations are not considered false positives due to quality assurance and quality control procedures.

In summary, it would appear that the presence and absence of viable enteric viruses can be successfully modeled in this watershed by using an approach that attempts to capture signals that distinguish fecal sources, indicate the approximate fecal age, and quantify the fecal load. The AC/TC ratio was the key to these modeling efforts, dominating the model and showing itself to be highly related to viral presence. We recommend that more researchers add the readily measured AC/TC ratio to their studies and that watershed management control strategies be implemented with AC/TC ratios as control or alert standards. Fecal sterols, with better methods that improve their recovery from environmental samples, should be investigated further for their utility in modeling the fecal loads of watersheds and identifying fecal sources.

ACKNOWLEDGMENTS

REFERENCES

1. **American Public Health Association.** 1992. Standard methods for the examination of water and wastewater, 18th ed. American Public Health Association, Washington, DC.
2. **Atherholt, T., E. Feerst, B. Hovendon, J. Kwak, and J. D. Rosen.** 2003. Evaluation of indicators of fecal contamination in groundwater. J. Am. Water Works Assoc. **95**(10):119–131.
3. **Booth, J., and G. M. Brion.** 2004. The utility of the AC/TC ratio for watershed management: a case study. Water Sci. Technol. **50**(1):199–203.
4. **Bosch, A.** 1998. Human enteric viruses in the water environment: a mini review. Int. Microbiol. **1**:191–198.
5. **Brion, G. M., J. S. Meschke, and M. D. Sobsey.** 2002. Male-specific coliphage: prevalence, types, and survival in natural waters. Water Res. **36**:2419–2425.
6. **Brion, G. M., and H. H. Mao.** 2000. Use of total coliform test for watershed monitoring with respect to atypicals. ASCE J. Environ. Eng. **126**(2):175–181.
7. **Brion, G. M., T. R. Neelakantan, and S. Lingireddy.** 2002. A neural network based classification scheme for sorting sources and ages of fecal contamination in water. Water Res. **36**:3765–3774.
8. **Cukor, G., and N. R. Blacklow.** 1984. Human viral gastroenteritis. Microbiol. Rev. **48**(2):157–179.
9. **Croci, L., M. D. De, C. Scalfaro, A. Fiore, M. Divizia, D. Donia, A. M. Cosentino, P. Moretti, and G. Costantini.** 2000. Determination of enteroviruses, hepatitis A virus, bacteriophages, and *Escherichia coli* in Adriatic Sea mussels. J. Appl. Microbiol. **88**:293–298.
10. **Dahling, D. R., and B. A. Wright.** 1986. Optimization of the BGM cell line culture and viral assay procedures for monitoring viruses in the environment. Appl. Environ. Microbiol. **51**:790–812.
11. **Deere, D., M. Stevens, A. Davison, G. Helm, and A. Dufour.** 2001. Management strategies, p. 257–288. *In* L. Fewtrell and J. Bartram (ed.), Water quality: guidelines, standards for health, assessment of risk and risk management for water-related infectious disease. IWA Publishing on behalf of the World Health Organization, London, United Kingdom.
12. **Finch, G. R., and N. Fairbairn.** 1991. Comparative inactivation of poliovirus type 3 and MS2 coliphage in demand-free phosphate buffer by using ozone. Appl. Environ. Microbiol. **57**:3121–3126.
13. **Fout, G. S., B. C. Martinson, M. W. N. Moyer, and D. R. Dahling.** 2003. A multiplex reverse transcription-PCR method for detection of human enteric viruses in groundwater. Appl. Environ. Microbiol. **69**:3158–3164.
14. **Freitas, S. J., G. M. Brion, L. E. Black, and T. Coakley.** 2006. Predictive input parameters for enteric virus presence at the inlet of a potable water supply. Water Sci. Technol. **54**(3):17–21.
15. **Grabow, W. O. K.** 1968. The virology of wastewater treatment. Water Res. **2**:675–716.
16. **Havelaar, A. H., M. van Olphen, and Y. C. Drost.** 1993. F-specific RNA bacteriophages are adequate model organisms for enteric viruses in fresh water. Appl. Environ. Microbiol. **59**:2956–2962.
17. **Havelaar, A. H., and M. D. Sobsey.** 1995. Detection of fRNA coliphages in groundwater. Lett. Appl. Microbiol. **20**:396–397.
18. **Hurst, C. J., G. R. Knudsen, M J. McInerney, L. D. Stetzenbach, and M. V. Walter.** 2001. Manual of environmental microbiology. ASM Press, Washington, DC.
19. **Jofre, J., E. Olle, F. Ribas, A. Vidal, and F. Lucena.** 1995. Potential usefulness of bacteriophages that infect *Bacteroides fragilis* as model organisms for monitoring virus removal in drinking water treatment plants. Appl. Environ. Microbiol. **61**:3227–3231.
20. **Kopecka, H., S. Dubrou, J. Prevot, J. Marechal, and J. M. Lopez-Pila.** 1993. Detection of naturally occurring enteroviruses in waters by reverse transcription, polymerase chain reaction, and hybridization. Appl. Environ. Microbiol. **59**:1213–1219.
21. **Leeming, R., A. Ball, N. Ashbolt, and P. Nichols.** 1996. Using fecal sterols from humans and animals to distinguish fecal pollution in receiving waters. Water Res. **30**:2893–2900.
22. **Leeming, R., and P. D. Nichols.** 1996. Concentrations of coprostanol that correspond to existing bacterial indicator guideline limits. Water Res. **30**:2997–3006.
23. **Long, S. C., S. S. El-Khoury, S. Oudejans, M. D. Sobsey, and J. Vinje.** 2005. Assessment of sources and diversity of male-specific coliphages for source tracking. Environ. Eng. Sci. **22**:367–377.
24. **Long, S. C., and M. D. Sobsey.** 2004. A comparison of the survival of F+RNA and F+DNA coliphages in lake water microcosms. J. Water Health **2**:15–22.
25. **Macdonald, I. A., V. D. Bokkenheuser, J. Winter, A. M. McLernon, and E. H. Mosbach.** 1983. Degradation of fecal sterols in the human gut. J. Lipid Res. **24**:675–694.
26. **Martins, C., R. C. Montone, R. C. Gamba, and V. H. Pellizari.** 2005. Sterols and fecal indicator microorganisms in sediments from Admiralty Bay, Antarctica. Braz. J. Oceanogr. **53**:1–12.
27. **Melnick, J. L., R. Safferman, V. C. Rao, S. Goyal, G. Berg, D. R. Dahling, B. A. Wright, E. Akin, R. Stetler, C. Sorber, B. Moore, M. D. Sobsey, R. Moore, A. L. Lewis, and F. M. Wellings.** 1984. Round robin investigation on methods for the recovery of poliovirus from drinking water. Appl. Environ Microbiol. **47**:144–150.
28. **Melnick, J. L.** 1984. Enteric viruses in water, vol.15, p. 1–16. Karger, Basel, Switzerland.
29. **Moriñigo, M. A., D. Wheeler, C. Betty, C. Jones, M. A. Muñoz, R. Cornax, and J. J. Borrego.** 1992. Evaluation of different bacteriophage groups as fecal indicators in contaminated natural waters in southern England. Water Res. **26**:267–271.
30. **Neelakantan, T., G. M. Brion, and S. Lingireddy.** 2001. Neural network modeling of *Cryptosporidium* and *Giardia* concentrations in the Delaware River. Water Sci. Technol. **43**(12):125–132.
31. **Nieman, J., and G. M. Brion.** 2003. Novel bacterial ratio for predicting fecal age. Water Sci. Technol. **47**(3):45–49.
32. **Payment, P., A. Berte, M. Prevost, B. Ménard, and B. Barbeau.** 2000. Occurrence of pathogenic microorganisms in the Saint-Lawrence River (Canada) and comparison of health risks for populations using it as their source of drinking water. Can. J. Microbiol. **46**:565–576.
33. **U.S. Environmental Protection Agency.** 1986. Ambient water quality criteria for bacteria. EPA 440/5-84-002. U.S. Environmental Protection Agency, Washington, DC.
34. **U.S. Environmental Protection Agency.** 1996. Information collection rule (ICR) microbial laboratory manual. EPA/600/R-95/178. U.S. Environmental Protection Agency, Washington, DC.

35. **U.S. Environmental Protection Agency.** 2001. Manual of methods for virology. EPA-600/4-84/013. U.S. Environmental Protection Agency, Washington, DC.

36. **U.S. Environmental Protection Agency.** 2001. Method 1601: male-specific (F+) and somatic coliphage in water by two-step enrichment procedure. EPA 821-R-01-030. U.S. Environmental Protection Agency, Washington, DC.

37. **U.S. Environmental Protection Agency.** 2001. Method 1602: male-specific (F+) and somatic coliphage in water by single agar layer (SAL) procedure. EPA 821-R-01-029. U.S. Environmental Protection Agency, Washington, DC.

38. **U.S. Environmental Protection Agency.** 2005. Method 1623: *Cryptosporidium* and *Giardia* in water by filtration/IMS/FA. EPA-815-R-05-002. U.S. Environmental Protection Agency, Washington, DC.

39. **U.S. Geological Survey.** 14 May 1997, modification date. Programs in Kentucky. Fact sheet FS-016-96. U.S. Geological Survey, Washington, DC. http://pubs.usgs.gov/fs/FS-016-96/.

40. **Venkatesan, M. I., and C. A. Santiago.** 1989. Sterols in ocean sediments: novel tracers to examine habitats of cetaceans, pinnipeds, penguins and humans. Mar. Biol. **102:**432–437.