ESCHERICHIA COLI AND FECAL-COLIFORM BACTERIA AS INDICATORS OF RECREATIONAL WATER QUALITY

By Donna S. Francy, Donna N. Myers, and Kevin D. Metzker

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 93-4083

Prepared in cooperation with the

CITY OF COLUMBUS, DIVISION OF SEWERAGE AND DRAINAGE;

CITY OF AKRON, PUBLIC UTILITIES BUREAU;

SUMMIT COUNTY DEPARTMENT OF ENVIRONMENTAL SERVICES;

OHIO WATER DEVELOPMENT AUTHORITY;

NORTHEAST OHIO REGIONAL SEWER DISTRICT; AND

CUYAHOGA RIVER COMMUNITY PLANNING ORGANIZATION

Columbus, Ohio
1993

U.S. DEPARTMENT OF THE INTERIOR

BRUCE BABBIT, Secretary

U.S. GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information write to:

District Chief
Water Resources Division
U.S. Geological Survey
975 W. Third Avenue
Columbus, OH 43212-3192

Copies of this report can be be purchased from:

U.S. Geological Survey
Books and Open-File Reports Section
Box 25425
Denver Federal Center
Denver, CO 80225

CONTENTS

Page

ILLUSTRATIONS

ILLUSTRATIONS--Continued

Page

TABLES

ESCHERICHIA COLI AND FECAL-COLIFORM BACTERIA AS INDICATORS OF RECREATIONAL WATER QUALITY

By Donna S. Francy, Donna N. Myers, and Kevin D. Metzker

ABSTRACT

In 1986, the U.S. Environmental Protection Agency (USEPA) recommended that *Escherichia coli* (*E. coli*) be used in place of fecal-coliform bacteria in State recreational water-quality standards as an indicator of fecal contamination. This announcement followed an epidemiological study in which *E. coli* concentration was shown to be a better predictor of swimming-associated gastrointestinal illness than fecal-coliform concentration. Water-resource managers from Ohio have decided to collect information specific to their waters and decide whether to use *E. coli* or fecal-coliform bacteria as the basis for State recreational water-quality standards. If one indicator is a better predictor of recreational water quality than the other, and if the relation between the two indicators is variable, then the indicator providing the most accurate measure of recreational water quality should be used in water-quality standards.

Water-quality studies of the variability of concentrations of *E. coli* to fecal-coliform bacteria have shown that (1) concentrations of the two indicators are positively correlated, (2) *E. coli* to fecal-coliform ratios differ considerably from site to site, and (3) the *E. coli* criteria recommended by USEPA may be more difficult to meet than current (1992) fecal-coliform standards. In this study, a statistical analysis was done on concentrations of *E. coli* and fecal-coliform bacteria in water samples collected by two government agencies in Ohio—the U.S. Geological Survey (USGS) and the Ohio River Valley Water Sanitation Commission (ORSANCO). Data were organized initially into five data sets for statistical analysis: (1) Cuyahoga River, (2) Olentangy River, (3) Scioto River, (4) Ohio River at Anderson Ferry, and (5) Ohio River at Cincinnati Water Works and Tanners Creek. The USGS collected the data in sets 1, 2, and 3, whereas ORSANCO collected the data in sets 4 and 5.

The relation of *E. coli* to fecal-coliform concentration was investigated by use of linear-regression analysis and analysis of covariance. Log-transformed *E. coli* and fecal-coliform concentrations were highly correlated in all data sets (r-values ranged from 0.929 to 0.984). Linear regression analysis on USGS and ORSANCO data sets showed that concentration of *E. coli* could be predicted from fecal-coliform concentration (coefficients of determination (R^2) ranged from 0.863 to 0.970). Results of analysis of covariance (ANCOVA) indicated that the predictive equations among the three USGS data sets and two ORSANCO data sets were not significantly different and that the data could be pooled into two large data sets, one for USGS data and one for ORSANCO data. However, results of ANCOVA indicated that USGS and ORSANCO data could not be pooled into one large data set.

Predictions of *E. coli* concentrations calculated for USGS and ORSANCO regression relations, based on fecal-coliform concentrations set to equal Ohio water-quality standards, further showed the differences in *E. coli* to fecal-coliform relations among data sets. For USGS data, a predicted geometric mean of 176 col/100 mL (number of colonies per 100 milliliters) was greater than the current geometric-mean *E. coli* standard for bathing water of 126 col/100 mL. In contrast, for ORSANCO data, the predicted geometric mean of 101 col/100 mL was less than the current *E. coli* standard.

The risk of illness associated with predicted *E. coli* concentrations for USGS and ORSANCO data was evaluated by use of the USEPA regression equation that predicts swimming-related gastroenteritis rates from *E. coli* concentrations.[1] The predicted geometric-mean *E. coli* concentrations for bathing water of 176 col/100 mL for USGS data and 101 col/100 mL for ORSANCO data would allow 9.4 and 7.1 gastrointestinal illnesses per 1,000 swimmers, respectively. This prediction compares well with the illness rate of 8 individuals per 1,000 swimmers estimated by the USEPA for an *E. coli* concentration of 126 col/100 mL. Therefore, the predicted geometric-mean *E. coli* concentration for bathing water seems to indicate a similar level of risk regardless of whether USGS or ORSANCO data are used.

Athough *E. coli* concentrations correlated well with fecal-coliform concentrations, the statistical relations between *E. coli* and fecal-coliform concentrations in natural waters can differ from one source of data to another. The epidemiological literature showed the relation between concentrations of *E. coli* and rate of swimming-associated illness was strong and consistent over geographic boundaries, whereas the relation between fecal-coliform bacteria and swimming-associated illness was not. Therefore, the difference between the use of *E. coli* and fecal-coliform bacteria is that *E. coli* can be used to establish standards based on an acceptable level of risk of swimming-associated illness, whereas fecal-coliform bacteria cannot.

INTRODUCTION

In 1986, the U.S. Environmental Protection Agency (USEPA) recommended that *Escherichia coli* (*E. coli*) or enterococci be used in place of fecal-coliform bacteria in State water-quality standards for the protection of people engaged in water-contact recreation (U.S. Environmental Protection Agency, 1986a). This recommendation was based on the results of a USEPA study (Dufour, 1984) in which a statistically significant relation was found between the rate of swimming-associated gastrointestinal illness and the concentration of *E. coli* and enterococci at freshwater beaches. The same study and a study done at marine locations (Cabelli, 1981) found no statistical relation between fecal-coliform concentration and swimming-associated gastrointestinal illness.

---

[1] Dufour, A.P., 1984, Health effects criteria for fresh recreational waters: Cincinnati, Ohio, U.S. Environmental Protection Agency, EPA-600/1-84-004, 33 p.

Because of the strong relation between *E. coli* concentration and gastrointestinal illness rate, the USEPA recommended *E. coli* criteria that are designed to provide the same level of protection afforded by the currently used State fecal-coliform standards (U.S. Environmental Protection Agency, 1986a). Criteria are elements of water-quality standards, expressed as constituent concentrations, levels, or narrative statements, representing a quality of water that supports a particular designated use. Water-quality standards are the rules set forth by the State establishing stream-use designations and water-quality criteria to protect users of State waters (Ohio Environmental Protection Agency, 1990). The *E. coli* criteria recommended by the USEPA, however, are considerably lower than the fecal-coliform standards for the same water-use designations in Ohio and several other states. In addition, the results of water-quality studies (Fandrei, 1985; Gannon and Busse, 1989; Milligan, 1987) indicate that, for some waters in the United States, the recommended *E. coli* criteria provide recreational users with a greater level of protection than do the established State fecal-coliform standards.

E. coli is part of the fecal-coliform group and has been credited to be a more specific indicator of fecal contamination than the more general test for fecal-coliform bacteria. Indeed, the presence of non-fecal-coliform bacteria from sources other than the gastrointestinal tracts of humans and other warm-blooded animals, which are thermotolerant and grow on fecal-coliform plates, can reduce the usefulness of fecal coliforms as an indicator of fecal contamination of surface water (Dufour, 1984; Dufour and Cabelli, 1976; Campbell and others, 1976; Caplenas and others, 1981).

The USEPA recommends the States adopt *E. coli* criteria as the basis for State recreational water-quality standards as soon as practicable. Water-resources managers in Ohio have chosen to collect information specific to State waters and decide whether to use *E. coli* or fecal-coliform bacteria as the basis for Ohio's recreational standards. The decision to retain either *E. coli* or fecal-coliform bacteria is to be based solely on the ability of the indicator bacteria to predict swimming-associated illness. The State of Ohio has temporarily adopted both *E. coli* and fecal-coliform criteria into State water-quality standards (Ohio Environmental Protection Agency, 1990). If the presence of *E. coli* in Ohio waters can be predicted consistently from fecal-coliform concentration, the inclusion of both indicator organisms in State water-quality standards may be unnecessary.

In this study, a statistical analysis was done on concentrations of *E. coli* and fecal-coliform bacteria in water samples collected by two government agencies. Data collected for this study were obtained in part from the Ohio River Valley Water Sanitation Commission (ORSANCO), and from cooperative projects of the U.S. Geological Survey (USGS) with the City of Columbus, Division of Sewerage and Drainage; the City of Akron, Public Utilities Bureau; Summit County Department of Environmental Services; Ohio Water Development Authority; the Northeast Ohio Regional Sewer District; and the Cuyahoga River Community Planning Organization.

Purpose and Scope

This report provides information on the use of *E. coli* and fecal-coliform bacteria as indicators of recreational water quality, not only to enhance the science and technical understanding of bacteria as water-quality indicators, but also to aid policy makers in Ohio in deciding whether to use

E. coli or fecal-coliform bacteria as the preferred indicator of recreational water quality on which to base State standards. The report includes a literature review on (1) the development of recreational water-quality criteria, (2) the link between recreational water quality and the occurrence of swimming-associated illness, and (3) comparisons of *E. coli* and fecal-coliform concentrations and their relation to each other and to State water-quality standards. The report also presents the results of a study in which bacteriological water-quality data were collected from four Ohio rivers (Scioto, Olentangy, Cuyahoga, and Ohio Rivers) to determine if a statistical relation exists between *E. coli* and fecal-coliform concentrations. The statistical relation was further examined to determine if *E. coli* concentration could be predicted accurately from fecal-coliform-concentration data and to what extent Ohio water-quality standards were exceeded by use of the two indicators.

REVIEW OF PREVIOUS INVESTIGATIONS

Two types of fecal-indicator-bacteria studies are discussed—epidemiological studies and water-quality studies. Epidemiological studies provide information for assessing the usefulness of concentrations of *E. coli* and fecal-coliform bacteria as predictors of swimming-associated illness. Water-quality studies provide information for assessing (1) the occurrence and ratios of *E. coli* and fecal-coliform concentrations in recreational waters outside Ohio and (2) the relation of concentrations of fecal-indicator bacteria to USEPA-recommended *E. coli* criteria and State water-quality standards.

Epidemiological Studies and the Development of Recreational Water-Quality Criteria

Since the 1930's, regulatory agencies have used concentrations of indicator organisms to estimate the health risk of recreational use of natural waters. For an indicator organism to be a reliable measure of water quality and health risk, however, the relation between concentration of indicator organism and rate of swimming-associated illness must be shown. Only recently was this relation firmly established (Cabelli, 1981; Dufour, 1984).

Studies following outbreaks of swimming-associated disease have long been used to describe the link between illness and fecal-contaminated waters. Although the agent causing the disease (etiological agent) was commonly identified, these studies lacked data describing water quality at the time of the swimmers' contact with the water. Similarly, early retrospective epidemiological studies, which involved interviews of disease-stricken individuals about their swimming habits, identified exposure factors associated with an etiological agent. These retrospective studies failed to show a direct relation between water quality and illness rate (Dufour, 1984). From about 1930 until 1968, many States used a total coliform-bacteria standard of 1,000 col/100 mL for recreational waters. This standard, however, was derived in a variety of ways and was commonly an arbitrary value that was easily attained and aesthetically acceptable (Dufour, 1984). Because these standards were derived by use of less-than-rigorous techniques according to today's standards, new studies were initiated.

In 1968, the National Technical Advisory Committee (NTAC) to the Federal Water Pollution Control Administration recommended bathing-water criteria on the basis of the results of three prospective epidemiological studies in which participants were selected before the onset of illness. Each study was done in a different geographic location during the late 1940's and early 1950's by the U.S. Public Health Service (USPHS) (Stevenson, 1953). Swimmers were asked to record swimming activity on calendars; total coliform bacteria concentrations also were monitored. Two out of three studies, both at freshwater locations, showed a significant increase in illness rate of swimmers who used the beaches when median total coliform concentrations were greater than 2,300 col/100 mL compared with swimmers who used the beaches when the water contained lower coliform concentrations. Because about 18 percent of total coliforms were found to be fecal coliforms, and given a reasonable margin of safety, the NTAC recommended bathing-water criteria based on a fecal-coliform concentration of 200 col/100 mL (U.S. Environmental Protection Agency, 1986b). In 1972, the USEPA adopted the same criteria and substantiated the use of this criteria primarily on the basis of known Salmonella/fecal-coliform ratios (Dufour, 1984). This criteria is still widely used as the basis for establishment of State water-quality standards today (Dufour, 1984).

The USPHS studies have been criticized over the years for several aspects of experimental design, including an imprecise definition of swimming practices and a statistically biased selection of data (U.S. Environmental Protection Agency, 1986b). Consequently, in 1972, the USEPA initiated a series of studies (Cabelli, 1981; Dufour, 1984) to determine the health risks of swimming in fecal-contaminated waters and to develop water-quality criteria for recreational waters. The investigators interviewed swimmers about rigidly defined swimming activities and well-defined disease symptoms. In addition, concentrations of fecal-coliform bacteria, *E. coli*, and enterococci were monitored throughout the studies in an effort to establish a direct link between swimming-associated gastroenteritis and concentrations of fecal-indicator bacteria.

The USEPA studies consisted of two phases: the first at several marine locations during 1972-1978 (Cabelli, 1981) and the second at two freshwater locations during 1978-82 (Dufour, 1984). Two paired beaches were selected for study at each marine and freshwater location: one where water quality was good and the other where water quality was barely acceptable with regard to local recreational-water-quality standards. At marine and freshwater locations, a statistically significant excess of gastrointestinal illness was found in swimmers who bathed at the barely acceptable beaches compared with those who bathed at the beaches with good water quality. In contrast, illness symptoms unrelated to gastroenteritis (respiratory, ear, nose, eye, and other ailments) were not excessive at any of the paired beaches (U.S. Environmental Protection Agency, 1986b).

Cabelli (1981) found a cause-effect relation between the concentration of enterococci and the rate of swimming-associated gastroenteritis at marine beaches. Concentration of fecal-coliform bacteria, the currently (1992) used indicator organism, did not correlate with the rate of illness. Similarly, Dufour (1984), in a freshwater study, found a strong correlation between the rate of swimming-associated gastroenteritis and the concentrations of *E. coli* and enterococci but not with fecal-coliform bacteria.

In 1986, as a result of the marine and freshwater studies, the USEPA recommended new recreational water-quality criteria based on a predictive model developed from data describing the relation between swimming-associated gastrointestinal illness and concentrations of *E. coli* or enterococci. The regression equation used by the USEPA to develop recreational water-quality criteria for *E. coli* in freshwater is

$$y = -11.74 + 9.40 (\log x), \qquad (1)$$

where y is the swimming-associated gastrointestinal illness rate per 1,000 swimmers, and x is the concentration of *E. coli* colonies per 100 mL (Dufour, 1984). The USEPA determined that a fecal-coliform concentration equal to the currently used geometric-mean standard for bathing waters of 200 col/100 mL would cause an estimated 8 illnesses per 1,000 swimmers at freshwater beaches. Substitution of 8 for y in equation 1 yields an *E. coli* concentration of 126 col/100 mL, the USEPA-recommended geometric-mean *E. coli* criteria for bathing waters.

The recommended geometric-mean *E. coli* criteria were calculated from the regression relation between number of indicator organisms and gastrointestinal illness rate to achieve the degree of protection provided with the currently used fecal-coliform standards (U.S. Environmental Protection Agency, 1986b). In addition, different confidence levels (c.l.) were used to calculate single-sample standards for four types of recreational use on the basis of the probability of full-body contact. The highest confidence levels correspond to the least restrictive conditions, as follows: infrequently used full-body-contact recreation, upper 95-percent c.l.; lightly used full-body-contact recreation, upper 90-percent c.l.; moderate full-body-contact recreation, upper 82-percent c.l.; and designated beach area, upper 75-percent c.l. In 1990, the state of Ohio set *E. coli* water-quality standards to accompany existing fecal-coliform standards for three types of recreational use—bathing waters, primary-contact waters, and secondary-contact waters (Ohio Environmental Protection Agency, 1990) (table 1).

Recent Epidemiological Studies

After the announcement of new USEPA recreational marine and freshwater criteria, other countries did regional epidemiological studies to determine links between swimming-associated illness and concentrations of bacteria (table 2).

In Canada, Seyfried and others (1985) studied the relation between bacterial concentration and swimmer-nonswimmer illness rates at several Ontario beaches where water quality was considered to be good. The investigators found a higher incidence of gastrointestinal illness among swimmers than among nonswimmers. A weak relation (correlation coefficient (r) = 0.284) (table 2) was found between fecal-coliform concentrations and the probability of contracting any swimming-related illness, including gastrointestinal illness; however, the relation between fecal-coliform concentrations and only gastrointestinal illness was not assessed. Although concentrations of other bacterial groups were monitored, *E. coli* concentrations were not assessed.

In a retrospective epidemiological study of a freshwater river in France during summer 1986 (Ferley and others, 1989), vacationers were interviewed about their swimming activities of the previous week and resulting illness type and duration. Concentrations of bacteria were also

Table 1.—Ohio numerical and narrative water-quality standards for bacteria in recreational waters

[Effective from May 1 through October 15. All values are in colonies per 100 milliliters. na, not applicable. Standards published in Ohio Environmental Protection Agency, 1990.]

Fecal-indicator bacteria type	Type of recreational water		
	Bathing[a] waters	Primary[b] contact	Secondary[c] contact
Fecal coliform:			
Geometric mean[d]	200	1,000	na
Single sample[e]	400	2,000	5,000
Escherichia coli:			
Geometric mean[d]	126	126	na
Single sample[e]	235	298	576

[a] Bathing waters are suitable for swimming and other full-body-contact exposure where a lifeguard or bathhouse is present.
[b] Primary-contact waters are suitable for full-body contact such as swimming, canoeing, and scuba diving.
[c] Secondary-contact waters are suitable for partial-body contact such as wading.
[d] The geometric mean is based on a minimum of five samples in a 30-day period.
[e] This value cannot be exceeded in more than 10 percent of the samples collected in a 30-day period.

monitored that week. Significantly fewer gastrointestinal illnesses were reported by swimmers on the "less polluted" beaches than on the "intermediate" and "more polluted" beaches. The investigators found poor correlation ($r = 0.38$), however, between gastrointestinal illness and fecal-coliform concentrations (table 2). Fecal-streptococci concentrations were better correlated with gastrointestinal illness ($r = 0.62$) than were fecal-coliform concentrations, whereas *E. coli* concentrations were not monitored.

The relation between concentrations of fecal indicators and swimming-associated illness was investigated at three beaches in Israel in 1983 (Fattal and others, 1987). Investigators considered two beaches to have "high fecal pollution" based on their proximities to sewage outfall; the third beach was considered to have "low fecal pollution" because it was not near any sewage outfall. The investigators found a higher illness rate in swimmers than in nonswimmers in the 0- to 4-year-old age group at the beaches having "high fecal pollution"; this difference in illness rate was not found for the higher age groups. In contrast, the investigators found no significant difference between the illness rate of swimmers and nonswimmers in any age group at the beaches having "low fecal pollution". In the 0- to 4-year-old age group at highly contaminated sites, high enterococci concentrations were most strongly associated with greater differences in the rates of gastrointestinal illness in swimmers compared to nonswimmers (level of significance, $p < 0.03$),

Table 2.—Summary of literature on relation between bacterial water quality and swimming-associated illness

[*E. coli*, *Escherichia coli*; ND, not determined; NS, not significant]

Aquatic setting and location (reference)	Gastrointestinal illness		General illness[e]	
	Fecal coliform	*E. coli*	Fecal coliform	*E. coli*
Freshwater lakes, United States (Dufour, 1984)	$r = 0.081$[a]	$r = 0.804$[a]	ND	ND
Marine beaches, United States (Cabelli, 1981)	$r = 0.01$[a]	$r = 0.512$[a]	ND	ND
Freshwater beaches, Canada (Seyfried and others, 1985)	ND	ND	$r = 0.284$[b]	ND
Freshwater rivers, France (Ferley and others, 1989)	$r = 0.38$[a]	ND	$r = 0.51$[a]	ND
Marine beaches, Israel (Fattal and others, 1987)	$p<0.08$[c]	$p<0.05$[c]	ND	ND
Marine beaches, Hong Kong (Holmes, 1989)	1.6($p<0.01$)[d]	1.5($p<0.05$)[d]	4.1($p<0.01$)[d]	4.0(NS)[d]

[a] Correlation coefficients for indicator density compared to the rate of illness.

[b] Correlation coefficients for indicator density compared to the odds, adjusted for significant factors, of swimmers becoming ill [$p/(1-p)$].

[c] Level of significance of difference in illness rate for swimmers in 0- to 4-year-old age group at sites with high numbers of bacterial indicators compared to sites with low numbers of bacterial indicators, based on one-way analysis of variance. A p-value less than or equal to 0.05 indicates that the difference in illness rates is significant.

[d] Difference in illness rate per 1,000 swimmers between swimming in "barely acceptable" and "relatively polluted" waters, as classified by numbers of the designated fecal indicator. The significance level is in parentheses. A p-value less than or equal to 0.05 indicates a significant difference between illness rate at barely acceptable compared to relatively polluted waters.

[e] A general illness refers to any swimming-associated illness, including, but not limited to, gastrointestinal illness.

high *E. coli* concentrations were somewhat associated with greater differences ($p < 0.05$), and high fecal-coliform concentrations were most weakly associated with greater differences ($p < 0.08$) (table 2).

The importance of recognizing regional water-quality and population differences was shown by a prospective epidemiological study of Hong Kong marine beaches in 1987 (Holmes, 1989). These beaches are relatively contaminated by Western standards; however, reported illness rates were less than those in Western studies. The investigators identified a moderate relation between *E. coli* concentration and the rate of swimming-associated gastroenteritis (table 2). In contrast to the results of other marine studies (Cabelli, 1981; Fattal and others, 1987), fecal-coliform concentrations were found to be linked to gastrointestinal and total illness rate, whereas enterococci concentration was not found to be a good predictor of swimming-associated gastroenteritis.

Comparisons of *Escherichia coli* and Fecal-Coliform Concentrations

In response to the USEPA recommendation to replace fecal coliforms with *E. coli* as indicators of recreational water quality, data submitted during the public-comment period showed that, at some beaches, a correlation was found between *E. coli* and fecal-coliform concentrations (U.S. Environmental Protection Agency, 1986b). Because there is a relation between *E. coli* concentration and the rate of swimming-associated gastroenteritis (Cabelli, 1981; Dufour, 1984), fecal-coliform concentration also may be related to gastroenteritis rate if fecal-coliform and *E. coli* concentrations are consistently correlated. Many studies, however, indicate a poor relation between fecal-coliform concentrations and the rate of swimming-associated gastroenteritis (Cabelli, 1981; Dufour, 1984; Ferley and others, 1989; Fattal and others, 1987). Therefore, observed ratios of ambient populations of *E. coli* and fecal-coliform bacteria in other studies may differ from ratios used in the work that led to the new (1986) *E. coli* criteria. Because the derivation of *E. coli* criteria was based on risk of illness associated with existing fecal-coliform standards, it is important to determine whether or not the *E. coli* to fecal-coliform (EC/FC) ratios found in recreational waters are similar to those promulgated in the standards and consistent across geographic boundaries.

Investigators monitored seven sites during wet- and dry-weather conditions for concentrations of fecal indicators in the Huron River near Ann Arbor, Mich. (Gannon and Busse, 1989). The EC/FC ratios, based on geometric means, ranged from 0.82 to 1.34 (table 3). Because these ratios were larger than 0.63, the ratio of the *E. coli* bathing-water criteria (126 col/100 mL) to the corresponding fecal-coliform standard (200 col/100 mL), the investigators concluded that the new *E. coli* criteria would be more difficult to meet than the current fecal-coliform standards.

Similarly, in a study of a recreational floatway in Alabama (Milligan, 1987), the investigators concluded that the water was rated to be more contaminated when new (1986) *E. coli* criteria were applied than when current fecal-coliform standards were used. In 4 out of the 13 months sampled, the monthly geometric mean exceeded the new *E. coli* criteria, whereas during only 1 month was the fecal-coliform standard exceeded. In all instances, the geometric-mean *E. coli* concentration was higher than the geometric-mean fecal-coliform concentration; EC/FC ratios ranged from 1.1 to 1.59 (table 3). The investigators found that fecal-coliform concentration was an excellent predictor of *E. coli* concentration for those waters (coefficient of determination (R^2) = 0.81).

Table 3.--Summary of literature on statistical relations between *Escherichia coli* and fecal-coliform concentrations

[EC, *Eschericia coli*; FC, fecal coliform; ND, not determined]

Study location (reference)	Land-use setting	EC/FC ratio	EC/FC regression statistics[a]	Regression equation
Michigan (Gannon and Busse, 1989)	Urban	0.82 - 1.34[b]	ND	ND
Alabama (Milligan, 1987)	Recreation and agriculture	1.1 - 1.59[b]	$R^2 = 0.81$	logEC = 0.88(logFC) + 0.73
Minnesota (Fandrei, 1985)	Urban	ND	$R^2 = 0.97$	lnEC = 0.95(lnFC) + 0.26
Israel (Fattal and others, 1987)	Coastal marine bathing beaches	0.283[c]	r = 0.88	ND

[a] R^2 is the coefficient of determination of the regression between log (ln) EC concentration and log (ln) FC concentration; r is the correlation coefficient.

[b] Based on the geometric mean.

[c] Based on mean log concentrations for 77 samples collected at 3 beaches.

In 1984, a study of the Mississippi and St. Croix Rivers in Minnesota was done to develop an understanding of fecal-indicator populations in local waters before, during, and after heavy rains (Fandrei, 1985). The investigator found that new *E. coli* criteria were exceeded more often and at more sampling locations than were currently used fecal-coliform standards. The author concluded that, for Minnesota waters, the proposed *E. coli* criteria seem to be more stringent than existing fecal-coliform standards. A strong relation was found between *E. coli* and fecal-coliform concentrations ($R^2 = 0.97$), especially after a heavy rain. The author suggested that this relation could be a result of fecal contamination from a common dominant source, such as combined-sewer overflows.

During an epidemiological study in Israel (Fattal and others, 1987), mean fecal-coliform concentrations for the entire 1983 bathing season were found to exceed mean *E. coli* concentrations significantly at three marine bathing beaches; the EC/FC ratio was 0.283. A high correlation was found between fecal-coliform and *E. coli* concentrations (table 3) (r = 0.88).

Therefore, in summary, the following can be said regarding fecal-coliform bacteria and *E. coli* as indicators of recreational water quality:

(1) The relation between the rate of swimming-associated gastroenteritis and fecal-coliform concentration is weak or absent in many cases, but a strong relation between swimming-associated gastroenteritis and *E. coli* concentration is fairly well documented;
(2) the concentrations of the two indicators are positively correlated;
(3) reported EC/FC ratios range from 0.283 to 1.59 compared to the EC/FC ratio of 0.63 in the bathing-water standards; and
(4) the proposed USEPA water-quality criteria based on *E. coli* could be more difficult to meet than existing fecal-coliform standards.

METHODS OF STUDY

The study examined the relation of *E. coli* to fecal-coliform concentrations in samples collected from State recreational waters. Sample-collection areas and sites, shown in figure 1, were the following:

(A) Columbus area—sites on the Olentangy and Scioto Rivers;
(B) Cleveland-Akron area—sites on the Cuyahoga River from Akron to the Cleveland navigation channel (including sites within the Cuyahoga Valley National Recreation Area) and the nearshore of Lake Erie at White City Beach in Cleveland; and
(C) the Cincinnati area—sites on the Ohio River.

Sample Collection and Analysis

Samples from 4 Olentangy River sites, 6 Scioto River sites, 11 sites along the Cuyahoga River, and 1 site near Cleveland in Lake Erie, were collected and analyzed by personnel of the Columbus, Ohio, office of the USGS. Sample results from three Ohio River sites were obtained from ORSANCO. The USGS and ORSANCO sites represent a range of primary-contact recreation and bathing waters in a variety of land-use settings including stream segments draining suburban, urban, and rural environments. Sites include discharger mixing zones, combined-sewer-overflow mixing zones, and mixed waters containing both discharger effluent and streamwater. U.S. Geological Survey samples were collected from inland rivers in Ohio, whereas ORSANCO samples were collected from a major river. All samples were collected during recreational seasons at base flow and during runoff-producing storms. In this report, ORSANCO data are used as an example of fecal-indicator data collected from Ohio waters and reported by another government agency, independent of the USGS.

Samples from the Olentangy and Scioto Rivers and 23 of the 31 samples from the Cuyahoga River were collected by use of flow-weighted sample-collection techniques (Ward and Harr, 1990). The other eight samples from the Cuyahoga River were grab samples. Grab samples collected for ORSANCO were obtained by volunteers at water-treatment-plant intakes or in mid-stream at the Anderson Ferry crossing. All samples were collected in sterile containers in a manner that minimized contamination. Samples were refrigerated at approximately 4°C and

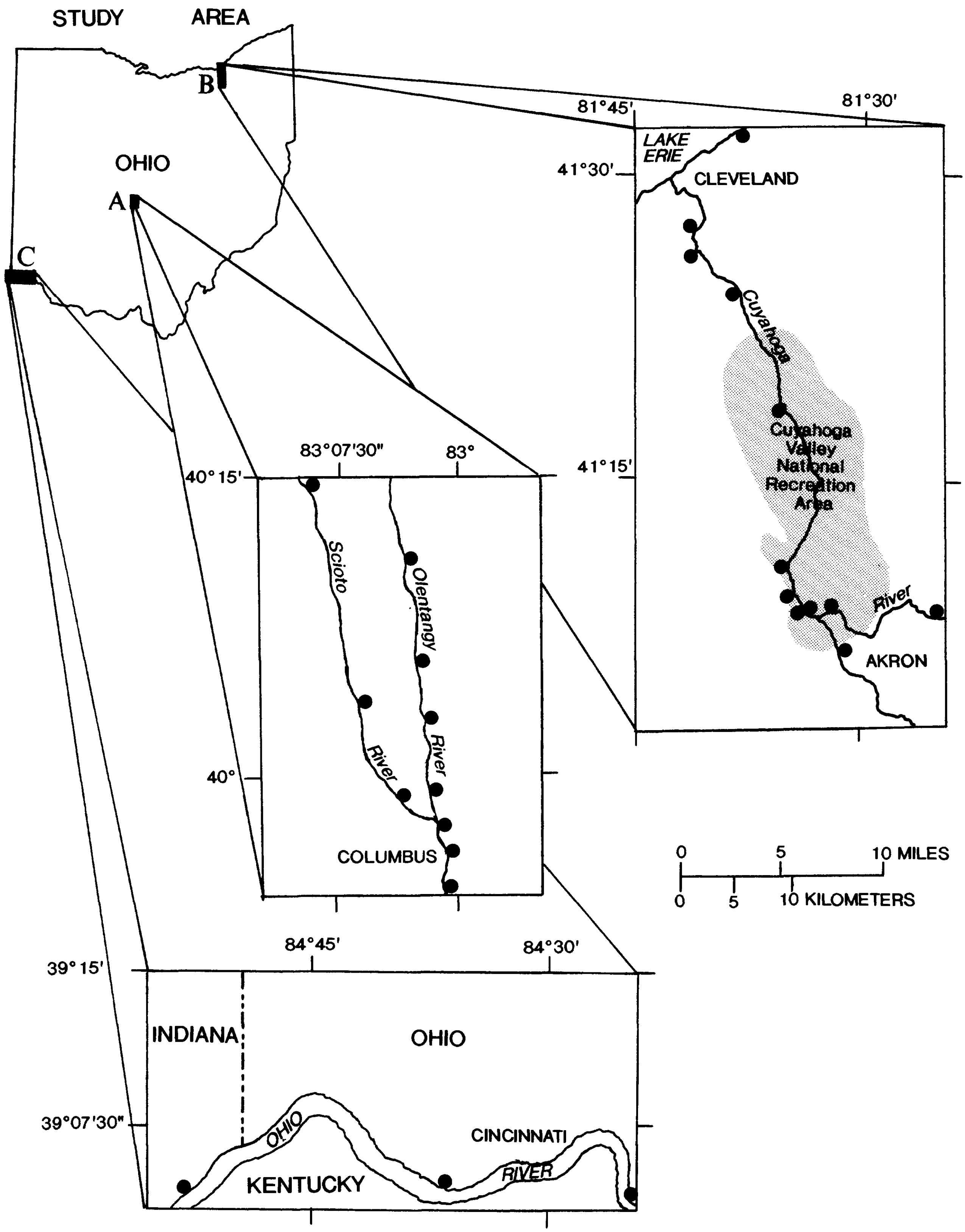


Figure 1.--Locations of sample-collection sites in (A) the Columbus, Ohio, area, (B) the Cleveland-Akron, Ohio, area, and (C) the Cincinnati, Ohio, area.

processed within 6 hours after collection at the USGS laboratory in Columbus, Ohio, and at the Fishcreek Wastewater Treatment Plant in Stow, Ohio (by USGS personnel); and, for ORSANCO, at a private laboratory in Cincinnati, Ohio.

E. coli and fecal-coliform bacteria were enumerated from aliquots taken from the same sample by use of standard membrane-filtration techniques. For *E. coli*, the m-TEC procedure (U.S. Environmental Protection Agency, 1985; American Public Health Association and others, 1989, sec. 9-50) was used for all samples. Samples collected by USGS were analyzed for fecal-coliform bacteria according to the procedure described by Britton and Greeson (1989); samples collected by ORSANCO were analyzed for fecal-coliform bacteria according to the standard method described by American Public Health Association and others (1989, sec. 9-94). Membrane filters with a pore size of 0.65 μm were used for samples collected from Olentangy and Scioto Rivers, whereas 0.45-μm-pore-size membrane filters were used for samples from Cuyahoga and Ohio Rivers. Although improved recoveries of fecal-coliform bacteria have been demonstrated by use of 0.65-μm-pore-size filters (Sladek and others, 1975; Lorenz and others, 1982), this bias was not important in this study because EC/FC data pairs were obtained by use of filters having the same pore size.

The fecal-coliform and *E. coli* methods (Britton and Greeson, 1989; U.S. Environmental Protection Agency, 1985) differ in several ways. The *E. coli* method includes a resuscitation step in which the bacteria are incubated for 2 hours at 35°C before incubation at 44.5°C for 22 to 24 hours. The resuscitation step allows stressed organisms to be recovered. In addition, the *E. coli* method includes a final step after incubation in which colonies are placed in a urea broth for 15 to 20 minutes. Only colonies remaining yellow, indicating a negative test for urease, are counted as *E. coli*.

Quality-assurance and quality-control practices were carried out through all phases of data collection and analysis by USGS and ORSANCO laboratory personnel. In the laboratory, quality-control testing of fecal-indicator methods and buffered water was done by use of coliform-reference samples obtained from the USEPA Quality Assurance Laboratory in Cincinnati, Ohio. Blanks prepared from 100 mL of dilution buffer were used as negative controls and filtered before and after each set of samples or at a frequency of no less than 10 percent. Laboratory practices—including specifications for dilution-water quality, cleaning practices, and safety precautions—were adopted as appropriate from guidelines set forth by Britton and Greeson (1989) and by Bordner and others (1978).

Although there may be variations in sampling or laboratory technique between agencies or different sites, these variations were not examined in this report. Relations between *E. coli* and fecal-coliform concentrations were examined, not absolute concentrations of bacteria. Standard methods, designed for consistency among users, were used by both USGS and ORSANCO.

Statistical Methods

Before statistical analysis, all data calculated from colony counts outside the ideal of 20 to 60 colonies per plate for fecal coliforms or 20 to 80 colonies per plate for *E. coli* were removed from the data sets. Paired concentrations of *E. coli* and fecal-coliform bacteria were statistically analyzed by use of the following data sets: (1-3) one for each of three streams sampled by the USGS (Scioto, Olentangy, and Cuyahoga Rivers); (4) one for the pooled USGS data; (5-6) two for data from three Ohio River sites sampled by ORSANCO; (5) Tanners Creek and Cincinnati Water Works, and (6) Anderson Ferry; and (7) one for the pooled ORSANCO data. Data set 5, from Cincinnati Water Works and Tanners Creek, is a combined set because few samples were collected at these sites, and because fecal-indicator concentrations of both sites were similar.

The statistical relation between *E. coli* and fecal-coliform concentrations was determined by linear-regression analyses of base 10 logarithmically transformed data. Regression diagnostics were done to test several assumptions of linear regression analysis that must be satisfied to provide the best linear, unbiased estimator of *E. coli* concentration from fecal-coliform concentration. The five assumptions requiring evaluation are that (1) the model form is correct (y is linear in x); (2) the data used to fit the model are representative of data which might be of interest; (3) the variance of the regression residuals is constant (it does not depend on the independent variable or other factors); (4) the regression residuals are independent; and (5) regression residuals are normally distributed.

Log transformations of fecal-indicator concentrations provided a linear fit of the data and satisfied the first assumption. The second assumption was met because the best direct methods available for analysis of fecal-indicator bacteria were used and because the waters from which samples were collected were designated for primary-contact or bathing recreation. To evaluate assumptions 3, 4, and 5, the investigators plotted regression residuals against observed log-transformed concentrations of fecal-coliform bacteria to determine whether residuals were similar in range and evenly distributed above and below the zero line over the entire range of observations. Regression residuals with substantial influence and leverage were evaluated. The regression residuals were tested to determine whether they were normally distributed by use of the Probability Plot Correlation Coefficient test (PPCC) (Looney and Gulledge, 1985).

Analysis of covariance (ANCOVA) was used to determine if the regression equations could be combined into pooled data sets for USGS data and ORSANCO data. The results of the ANCOVA dictated the manner in which to combine (pool) the data; data sets having statistically different slopes or y intercepts could not be pooled.

To determine what effect the relation beween fecal-coliform concentration and *E. coli* concentration would have on water-quality standards, the USGS and ORSANCO regression equations were used to predict *E. coli* concentrations from fecal-coliform concentrations when fecal-coliform concentrations were set to equal Ohio water-quality standards for bathing waters, primary-contact recreation, and secondary-contact recreation (Ohio Environmental Protection Agency, 1990) (table 1). Ninety-five-percent confidence intervals were calculated to estimate variability of mean *E. coli* concentrations about the regression line. In contrast, 90-percent prediction intervals estimate the variability of single *E. coli* concentrations about the regression line.

Confidence intervals of predicted concentrations are important for describing the range of probable values that the prediction can take on at a given level of uncertainty. If the true regression equation were known, then the predicted geometric-mean *E. coli* concentration (y), given a geometric-mean fecal-coliform concentration (x), would be determined from the relation

$$m_{(y/x)} = a + b(x).$$

However, a y intercept (a) and slope (b) are estimated by use of observed concentrations for x and y; hence, $m_{(y/x)}$ (mean of y given x) is estimated (Iman and Conover, 1983, p. 377). The range of probable values for a predicted geometric-mean *E. coli* concentration can be described by use of the confidence interval for $m_{(y/x)}$. The 95-percent confidence interval of the predicted geometric-mean *E. coli* concentration represents the range of values that the predicted geometric mean could assume in 95 of every 100 samples.

Prediction intervals for the geometric mean are wider than confidence intervals and, for purposes of this report, are used to describe the bounds of the predicted geometric-mean *E. coli* concentration (y) when a fecal-coliform concentration (x) from a single sample is the predictor (rather than the geometric-mean *E. coli* concentration). The upper 90-percent prediction limit was chosen because of the requirement in Ohio that no more than 1 in 10 samples collected within a 30-day period can exceed the single-sample standard. The upper 90-percent prediction limit of the predicted geometric-mean *E. coli* concentration represents the range of probable values that the geometric mean could assume in 9 of every 10 cases generated from single-sample predictions.

To determine whether *E. coli* and fecal-coliform standards were comparable as indicators of recreation-water quality, the investigators calculated the percentage of samples exceeding *E. coli* standards while meeting fecal-coliform standards for the same recreational use-designation.

RELATION OF *ESCHERICHIA COLI* CONCENTRATIONS TO FECAL-COLIFORM CONCENTRATIONS

The data representing the five unpooled data sets (1-3, 5, and 6) and the two pooled data sets (4 and 7) are shown on scatterplots (figs. 2 and 3); each point on the scatterplots of data represents a pair of log-transformed *E. coli* and fecal-coliform concentrations from a single sample. Log-transformed *E. coli* and fecal-coliform concentrations were highly correlated in all data sets; r-values ranged from 0.929 to 0.984 (table 4).

The standard error of the regression measures the degree of deviation of observed values from the regression line and is an indicator of the level of uncertainty associated with a prediction, expressed as a percentage of the predicted mean. The range of standard errors associated with the regression equations was 33.8 to 51.3 percent (table 4). Standard errors in this range indicate that predicted concentrations of *E. coli* can vary as little as 33.8 percent and as much as 51.3 percent.

The slope of the regression line is a measure of the rate of change in log *E. coli* with change in log fecal-coliform concentration. For some data sets, the rate of change of *E. coli* concentration with change in fecal-coliform concentration was nearly equal, whereas for other data sets, the rate of change was somewhat less than 1.0 (table 4). For USGS unpooled data, slopes ranged from

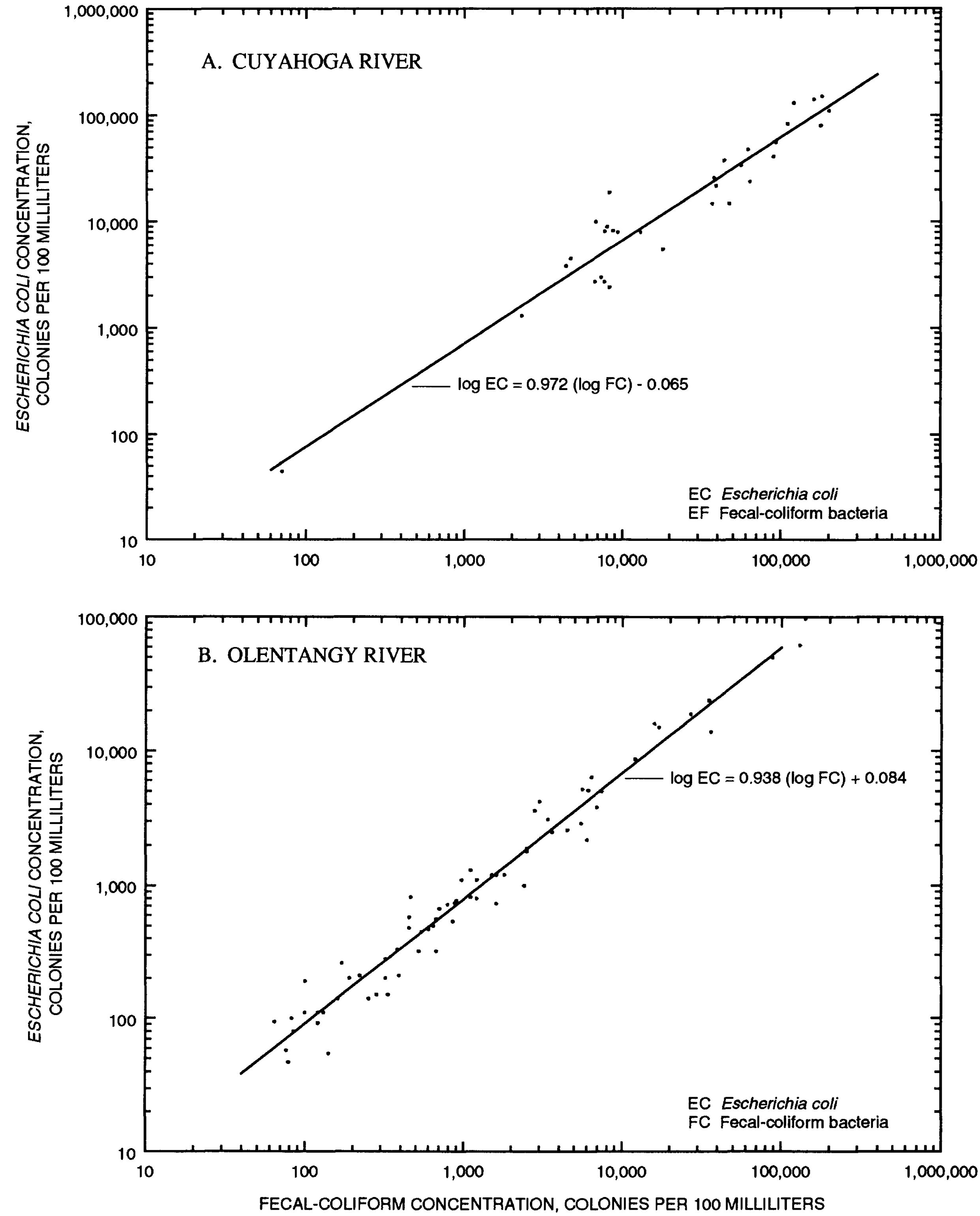
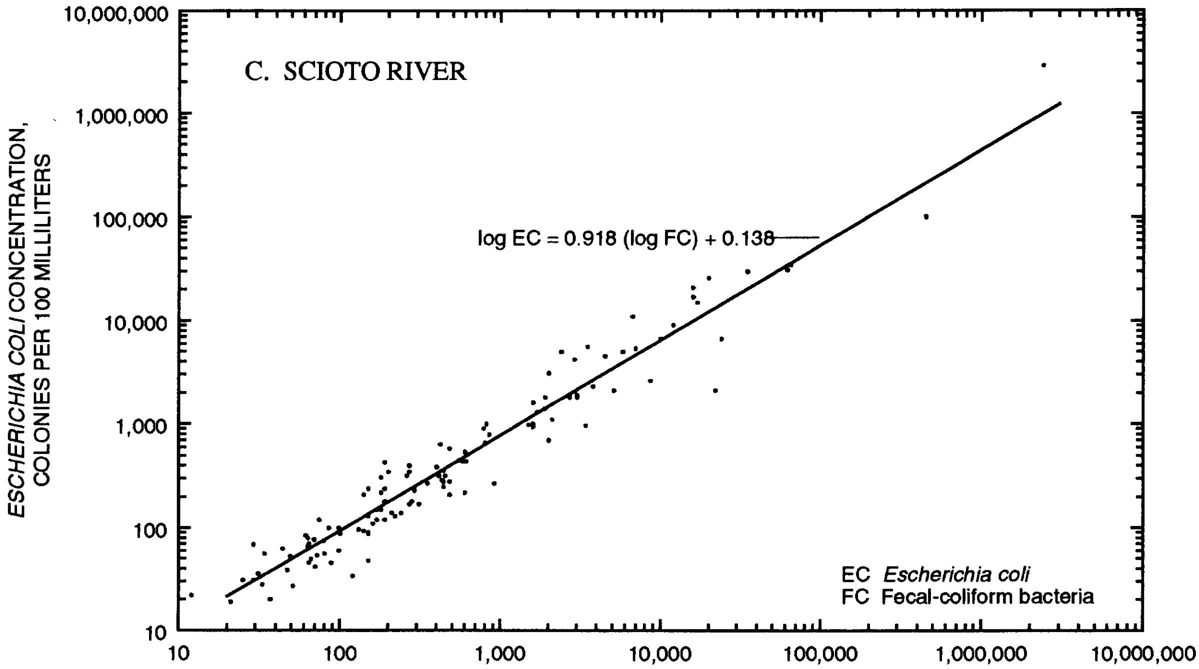


Figure 2.--Regression relation between *Escherichia coli* and fecal-coliform concentrations (A) in the Cuyahoga River, (B) in the Olentangy River, (C) in the Scioto River , and (D) for pooled U.S. Geological Survey data.



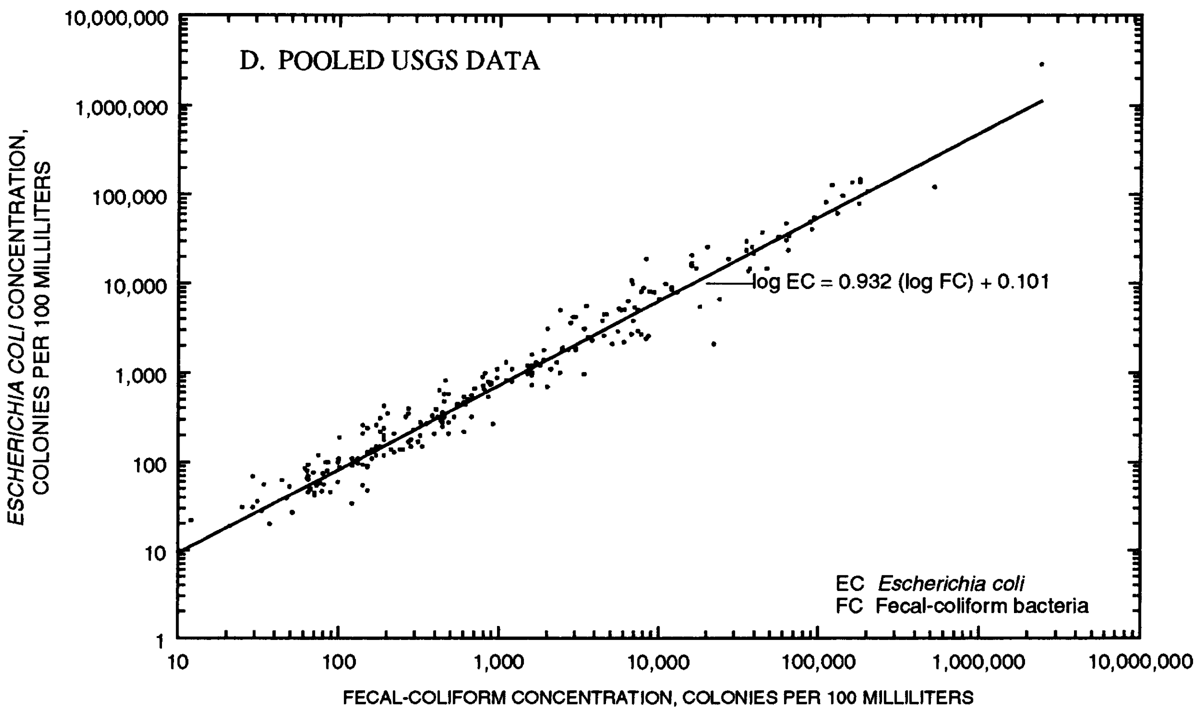


Figure 2.--Regression relation between *Escherichia coli* and fecal-coliform concentrations (A) in the Cuyahoga River, (B) in the Olentangy River, (C) in the Scioto River, and (D) for pooled U.S. Geological Survey data--Continued.

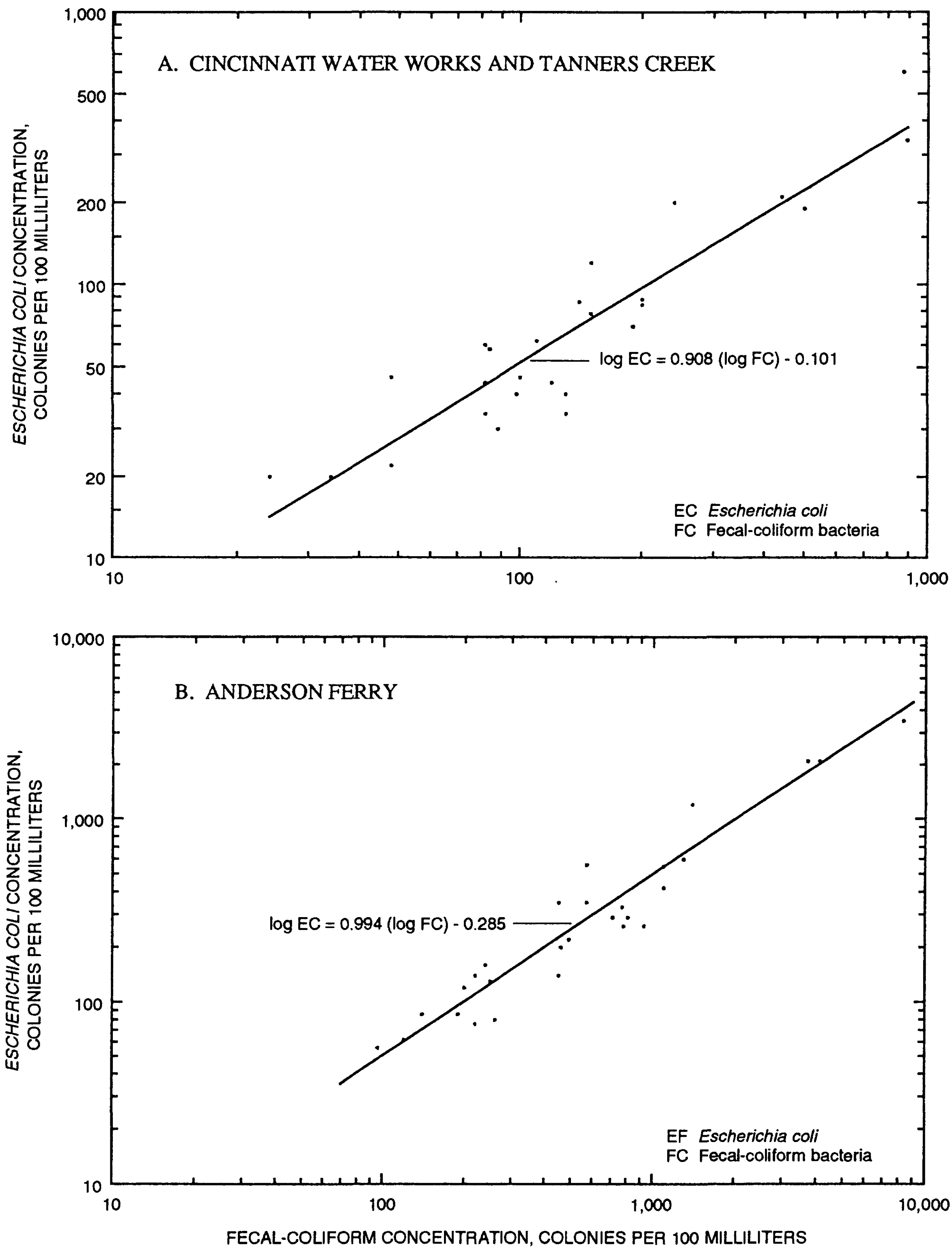


Figure 3.--Regression relation between *Escherichia coli* and fecal-coliform concentrations (A) at Cincinnati Water Works and Tanners Creek, (B) at Anderson Ferry, and (C) for pooled Ohio River Valley Water Sanitation Commission data.

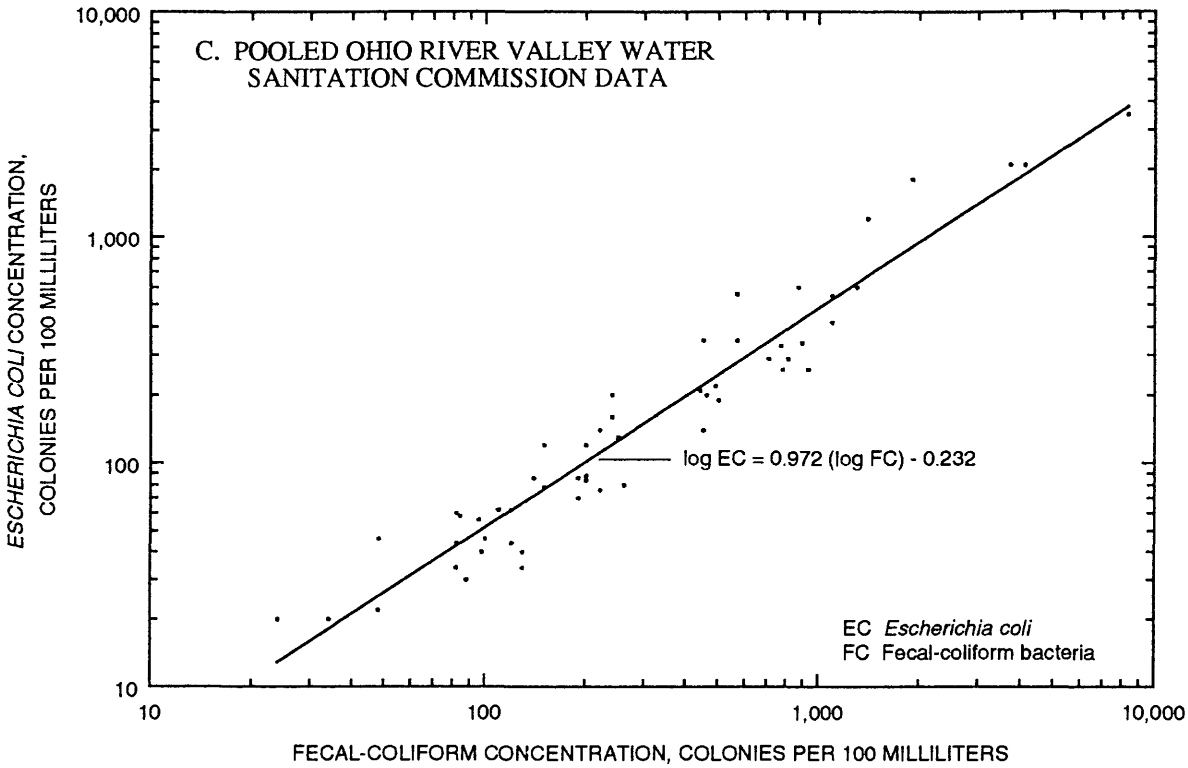


Figure 3.--Regression relation between *Escherichia coli* and fecal-coliform concentrations (A) at Cincinnati Water Works and Tanners Creek, (B) at Anderson Ferry, and (C) for pooled Ohio River Valley Water Sanitation Commission data--Continued.

0.918 for Scioto River data to 0.972 for Cuyahoga River data. For ORSANCO unpooled data, slopes were 0.908 for Cincinnati Water Works/Tanners Creek data and 0.994 for Anderson Ferry data. The *t*-tests ($\alpha = 0.05$) on the slopes of the regression lines from five unpooled data sets indicated that all five slopes were significantly different from zero at probabilities of less than 0.001 (table 4).

The *y* intercept is the value for log *E. coli* that corresponds to a zero value for log fecal-coliform bacteria. The *y* intercepts from five unpooled data sets ranged from -0.285 for Ohio River at Anderson Ferry to 0.138 for Scioto River (table 4). The *t*-tests ($\alpha = 0.05$) on the *y* intercepts indicated that four out of five were not significantly different from zero. Only the *y* intercept for Scioto River data was significantly different from zero (table 4).

The coefficient of determination (R^2) is the fraction of the variation in log *E. coli* concentration that is explained by log fecal-coliform concentration. For the five unpooled data sets, R^2 ranged from 0.863 to 0.970 (table 4), indicating that a large amount of the variation in *E. coli* concentration is explained by fecal-coliform concentration.

Table 4.--Regression statistics for *Escherichia coli* and fecal-coliform concentrations

The correlation coefficient (r) measures the strength of the linear association between *E. coli* and fecal-coliform concentrations. The standard error of the regression (S), also known as a standard deviation of the residuals, measures the degree of uncertainty associated with a prediction of concentrations of *E. coli* from fecal-coliform concentrations. The p-value of t-tests on the slopes of the regression lines is the probability that the null hypothesis (the slope is equal to zero) is true ($\alpha = 0.05$). The p-value of t-tests on the y intercepts of the regression lines is the probability that the null hypothesis (the y intercept is equal to zero) is true ($\alpha = 0.05$). The coefficient of determination (R^2) of the regression between log *E. coli* concentrations and log fecal-coliform concentrations is the fraction of the variation in *E. coli* concentration that can be explained by fecal-coliform concentration.

Data set	Sample size	r	S (percent)	Slope (log units)	p-value for t-tests on slope	y intercept (log units)	p-value for t-tests on y intercept	R^2
U.S. GEOLOGICAL SURVEY DATA								
Cuyahoga River	32	0.959	51.3	0.972	<0.001	-0.065	0.778	0.920
Olentangy River	69	.983	34.9	.938	<.001	.084	.227	.965
Scioto River	112	.977	40.2	.918	<.001	.138	.009	.954
Pooled data	213	.984	40.2	.932	<.001	.101	.006	.970
OHIO RIVER VALLEY WATER SANITATION COMMISSION (ORSANCO) DATA								
Anderson Ferry	29	0.956	34.9	0.994	<0.001	-0.285	0.095	0.913
Cincinnati Water Works/ Tanners Creek	27	.929	33.8	.908	<.001	-.101	.528	.863
Pooled data	56	.964	34.1	.972	<.001	-.232	.015	.929

Regression residuals, the differences between observed and predicted values for *E. coli* concentrations, were calculated for each of the five regression analyses. No curvature or bias was observed in scatterplots of regression residuals and observed values for four of the five unpooled data sets. One outlier in the Cincinnati Water Works/Tanners Creek data set caused the residuals to deviate from a normal distribution, owing to the high leverage and influence of the outlier.

The effect of this single outlier on the regression relation was further tested. Removal of the outlier from the data set increased the slope of the regression line from 0.757 to 0.908. Because the EC/FC ratio of this outlier was 0.085, which was considerably less than EC/FC ratios for the other 273 USGS and ORSANCO data pairs and considerably less than any published EC/FC ratio (table 3), the outlier was considered atypical and was removed from the data set. The values listed in table 4 for the Cincinnati Water Works/Tanners Creek data set were calculated after removal of this outlier. The PPCC test indicated that normality could not be rejected for the five unpooled data sets of regression residuals ($\alpha = 0.05$).

ANCOVA was used to determine if the regression equations from three USGS data sets and two ORSANCO data sets were statistically different with respect to slopes and *y* intercepts. If the regression equations were not statistically different, data could be combined into two pooled data sets, one for USGS data and one for ORSANCO data. The *F*-tests from ANCOVA compared the error-sum-of-squares values and residual degrees of freedom for each of two regression models: (1) a simple model of pooled USGS or ORSANCO data and (2) a complex model composed of two ORSANCO or three USGS unpooled sets in which additional slope and intercept terms were added by use of dummy variables. For pooled compared to unpooled sets of USGS or ORSANCO data, the calculated *F*-values were greater than the values in the *F*-distribution table at $\alpha = 0.05$. Thus, in both cases, the null hypothesis (that the coefficients of the additional variables in the complex model were not different from zero) was rejected, and the simple model of pooled data was chosen over the complex model. The pooled USGS data and pooled ORSANCO data were superior to unpooled data for explaining the relation between the two fecal-indicator bacteria. In addition, *t*-tests indicated that the *y* intercepts and slopes from each of the three USGS data sets and two ORSANCO data sets were not significantly different from each other (table 5). Therefore, data were pooled into two sets, one for USGS data and one for ORSANCO data.

Additional ANCOVA was done to determine whether USGS and ORSANCO data could be pooled into one data set. The calculated *F*-value was less than the *F*-value in the table: the *F*-test was significant at $\alpha = 0.05$, and the complex model (unpooled USGS and ORSANCO data) was chosen over the simple model (pooled USGS and ORSANCO data). The relation of *E. coli* to fecal-coliform concentration was different between the pooled data sets from each organization ($p < 0.001$) (table 5). The *y* intercepts and slopes of the regression lines on data collected by the two agencies were significantly different (fig. 4). Therefore, data were kept in two pooled data sets—a USGS data set and an ORSANCO data set—for further analysis.

Regression statistics describing the two pooled data sets show that the slope for pooled USGS data was 0.932 and for pooled ORSANCO data was 0.972 (table 4). Therefore, the rate of change in *E. coli* concentration based on fecal-coliform concentration was somewhat greater for ORSANCO data than for USGS data. The standard error was slightly smaller for the pooled

Table 5.--Results of analysis of covariance for pooled U.S. Geological Survey and Ohio River Valley Water Sanitation Commission regressions relating *Escherichia coli* and fecal-coliform concentrations

[USGS, U.S. Geological Survey; ORSANCO, Ohio River Valley Water Sanitation Commission; na, not applicable (refer to table 4 for specific information]

Data set	*F*-tests[a]		Model chosen	*t*-tests for coefficients[b] of model chose			
				Slope		Intercept	
	F-value	*p*-value		*t*-value	*p*-value	*t*-value	*p*-value
Pooled USGS	1.50	0.216	Simple (Pooled)	82.7	<0.001	2.84	0.005
Pooled ORSANCO	.531	>.25	Simple (Pooled)	26.61	<.001	-2.53	.015
USGS and ORSANCO	45.73	<.001	Complex (Unpooled)	na	na	na	na

[a] *F*-test from ANCOVA comparing regressions for simple model and complex model.
[b] *t*-test results for slope and *y* intercept coefficients of model chosen.

ORSANCO data (34.1 percent) than for the pooled USGS data (40.2 percent). The R^2 values were 0.970 for pooled USGS data and 0.929 for pooled ORSANCO data, indicating that greater than 90 percent of the variation in *E. coli* concentration was explained by fecal-coliform concentration. Of all the regression terms, the *y* intercepts were the most different between the two predictive equations and significantly affected the ratios of the two indicators. The *y* intercept for pooled USGS data was 0.101, whereas the *y* intercept for pooled ORSANCO data was -0.232 (fig. 4 and table 4).

PREDICTIONS OF *ESCHERICHIA COLI* FROM FECAL-COLIFORM CONCENTRATIONS AND THE RELATION OF THESE CONCENTRATIONS TO OHIO WATER-QUALITY STANDARDS

Regression equations for USGS and ORSANCO data were used to predict *E. coli* concentrations from different fecal-coliform concentrations to illustrate the variability of predicted *E. coli* concentrations resulting from the equation chosen. Predictions of *E. coli* concentrations made for this report are provided for interpretive purposes only and are not meant to replace current numerical standards or criteria.

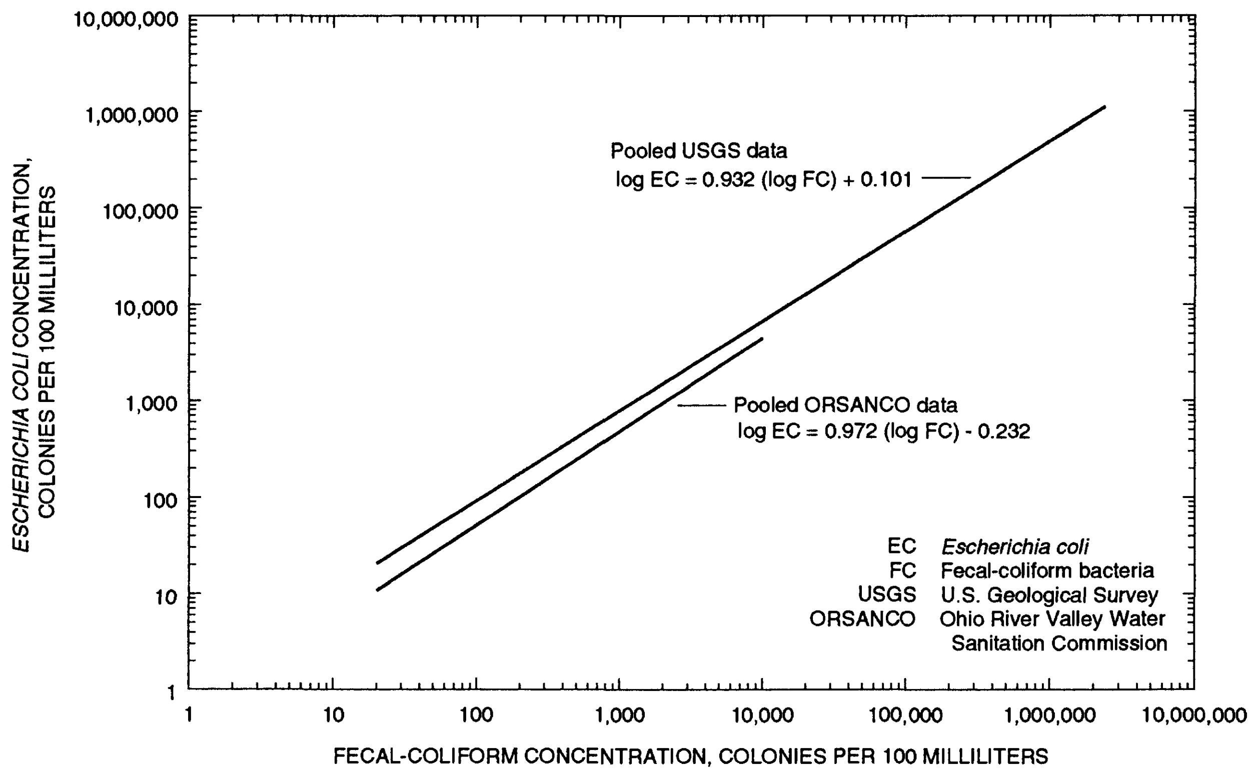


Figure 4.--Regression relations between *Escherichia coli* and fecal-coliform concentrations for U.S. Geological Survey and Ohio River Valley Water Sanitation Commission data.

E. coli concentrations predicted from regression equations for USGS and ORSANCO data are geometric means because they represent the antilog of the mean of log-transformed concentrations. Predicted geometric-mean *E. coli* concentrations and corresponding 95-percent confidence intervals and upper 90-percent prediction limits were calculated from USGS and ORSANCO regression relations by setting fecal-coliform concentrations to the geometric-mean Ohio water-quality standards for bathing water and primary-contact recreation (table 6). The antilogs of log-transformed *E. coli* concentrations predicted from log fecal-coliform concentrations, as well as confidence and prediction limits, are plotted for USGS and ORSANCO pooled data sets on the basis of each predictive equation (figs. 5 and 6).

For pooled USGS data, a geometric-mean *E. coli* concentration of 176 col/100 mL and a 95-percent confidence interval of 167 to 189 col/100 mL were predicted from a fecal-coliform concentration set to the geometric-mean bathing-water standard of 200 col/100 mL (table 6). A geometric-mean *E. coli* concentration of 790 col/100 mL and a 95-percent confidence interval of 759 to 841 col/100 mL were predicted from a fecal-coliform concentration of 1,000 col/100 mL,

Table 6.--Confidence and prediction intervals for geometric-mean *Escherichia coli* concentrations from regression relations with fecal-coliform concentrations set to equal Ohio water-quality standards

[*E. coli*, *Escherichia coli*; USGS, U.S. Geological Survey; ORSANCO, Ohio River Valley Water Sanitation Commission. All units are in col/100 mL, colonies per 100 milliliters]

Pooled data set	Fecal coliform standard	*E. coli* standard		Predicted *E. coli* geometric mean	95-percent confidence interval of geometric mean	Upper 90-percent prediction value of geometric mean	Ratio of predicted *E. coli* to fecal-coliform concentration
		Geometric mean	Single sample				
	200[a]	126[a]	235[d]	176	167 to 189	320	0.89
	1,000[b]	126[b]	298[e]	790	759 to 841	1,430	0.80
	5,000[c]		576[c]	3,540	--	6,420	0.72
ORSANCO	200[a]	126[a]	235[d]	101	92 to 111	170	0.51
	1,000[b]	126[b]	298[e]	484	426 to 549	816	0.48
	5,000[c]		576[c]	2,310	--	3,980	0.46

[a] Geometric-mean standard for bathing water.
[b] Geometric-mean standard for primary-contact recreation.
[c] Single-sample standard for secondary-contact recreation.
[d] Single-sample standard for bathing water.
[e] Single-sample standard for primary-contact recreation.

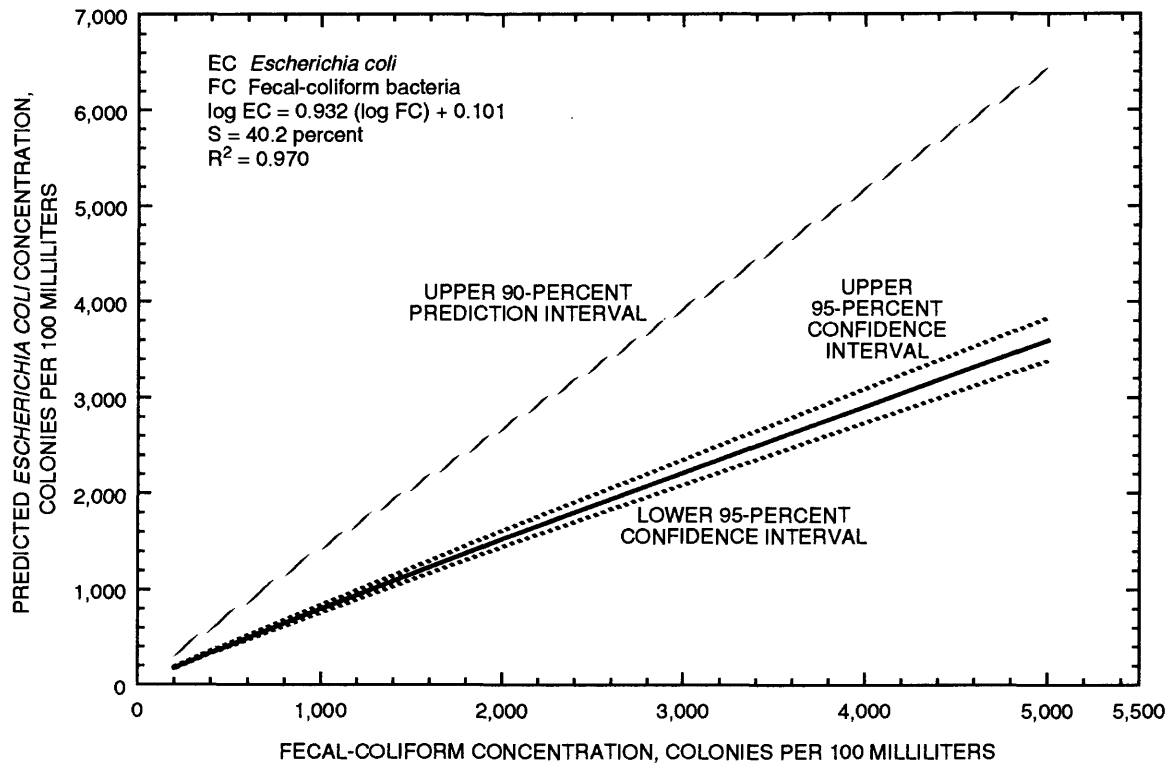


Figure 5.--Confidence and prediction intervals for *Escherichia coli* concentrations calculated from the fecal-coliform/*Escherichia coli* regression relations for pooled U.S. Geological Survey data. (Statistical terms are defined in table 4.)

the geometric-mean primary-contact standard. The current geometric-mean bathing-water and primary-contact standard of 126 col/100 mL is well below the predicted geometric mean and 95-percent confidence intervals for both recreational-use designations. In addition, the ratio of predicted geometric-mean *E. coli* to geometric-mean fecal-coliform concentration calculated from pooled USGS data for bathing water was 0.89 and for primary-contact water was 0.80 (table 6). These ratios are much higher than the EC/FC ratios of 0.63 for bathing water and 0.126 for primary-contact recreation for the current numerical water-quality standards.

For pooled USGS data, the *E. coli* single-sample bathing-water standard of 235 col/100 mL and single-sample primary-contact standard of 298 col/100 mL are well below the upper 90-percent prediction limits of 320 col/100 mL for bathing water and 1,430 col/100 mL for primary-contact recreation. Therefore, the single-sample water-quality standards will probably be exceeded in many more than 10 percent of samples because the current standards fall well below the upper limits of the 90-percent prediction intervals.

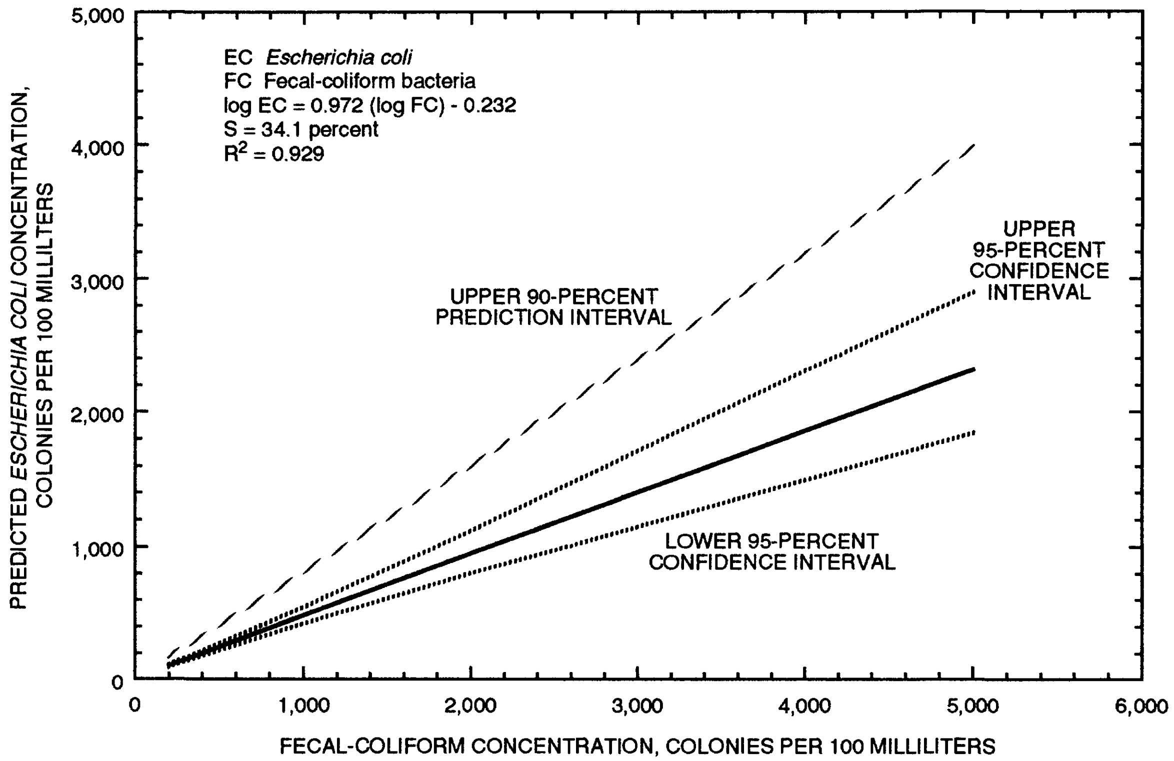


Figure 6.--Confidence and prediction intervals for *Escherichia coli* concentrations calculated from the fecal-coliform/*Escherichia coli* regression relations for pooled Ohio River Valley Water Sanitation Commission data. (Statistical terms are defined in table 4.)

For pooled ORSANCO data, a geometric-mean *E. coli* concentration of 101 col/100 mL and a 95-percent confidence interval of 92 to 111 col/100 mL were predicted from a fecal-coliform concentration set to the geometric-mean bathing-water standard of 200 col/100 mL. A geometric-mean *E. coli* concentration of 484 col/100 mL and a 95-percent confidence interval of 426 to 549 col/100 mL were predicted from a geometric-mean fecal-coliform concentration set to 1,000 col/100 mL, the geometric-mean primary-contact standard. Unlike the USGS results, the current geometric-mean bathing-water standard of 126 col/100 mL is greater than the predicted geometric-mean *E. coli* concentration and is above the 95-percent confidence interval of this predicted value. The ratio of predicted geometric-mean *E. coli* to geometric-mean fecal-coliform concentration calculated from pooled ORSANCO data for bathing water was 0.51, lower than the 0.63 ratio for current bathing-water standards. For primary contact, the 0.48 ratio of predicted geometric-mean *E. coli* concentration to geometric-mean fecal-coliform concentration is higher than the 0.126 EC/FC ratio for geometric-mean primary-contact recreational standards.

For pooled ORSANCO data, the single-sample bathing-water standard of 235 col/100 mL is greater than the upper 90-percent prediction limit of 170 col/100 mL. Therefore, the single-sample water-quality standard for bathing water will probably not be exceeded in more than 10 percent of all samples because the standard is greater than the concentration associated with the upper 90-percent prediction limit. The single-sample primary-contact standard of 298 col/100 mL, however, is well below the upper 90-percent prediction limit of 816 col/100 mL. Therefore, the single-sample primary-contact standard probably will be exceeded at rates much higher than the permissible rate of 10 percent of samples because the standard is much lower than the concentration associated with the upper 90-percent prediction limit.

To evaluate the statistical models in terms of actual data, the investigators calculated the percentage of samples that exceeded Ohio *E. coli* standards while attaining fecal-coliform standards for pooled USGS data sets and pooled ORSANCO data sets (table 7). For pooled USGS data, geometric-mean and single-sample standards were exceeded in all categories of recreational use more frequently for *E. coli* than for fecal-coliform bacteria. For bathing water, the geometric-mean standard for *E. coli* was exceeded in 23.9 percent of all samples when fecal-coliform concentrations from the same samples met the standard. The single-sample bathing-water standard for *E. coli* was exceeded in 16.1 percent of samples in which fecal-coliform concentrations met the standard. For primary-contact recreation, the geometric-mean standard for *E. coli* was exceeded in 58.5 percent of samples in which fecal-coliform concentrations from the same samples met the standard (table 7 and fig. 7). The single-sample primary-contact standard for *E. coli* was exceeded in 39.0 percent of samples in which fecal-coliform concentrations met the standard.

For pooled ORSANCO data, the *E. coli* water-quality standards for bathing waters were met in all samples for which the fecal-coliform standards were met (fig. 8). The geometric-mean primary-contact standard for *E. coli* was exceeded in 25.4 percent of all samples in which the fecal-coliform standard was met (table 7 and fig. 8). The single-sample primary-contact standard for *E. coli* was exceeded in 13.3 percent of all samples in which the fecal-coliform standards were met. For pooled ORSANCO data, the primary-contact standards for *E. coli* were met more than twice as often as for USGS data.

RISK OF ILLNESS BASED ON PREDICTIONS OF *ESCHERICHIA COLI* CONCENTRATIONS

To determine the risk of illness associated with predicted *E. coli* concentrations, the investigators substituted the predicted *E. coli* concentrations into the USEPA regession equation derived by Dufour (1984) (eq. 1). The predicted geometric-mean *E. coli* concentration for bathing waters of 176 col/100 mL for pooled USGS data and 101 col/100 mL for pooled ORSANCO data (table 6) substituted into equation 1 would generate 9.4 and 7.1 swimming-associated gastrointestinal illnesses per 1,000 swimmers, respectively. The illness rate of 8 individuals per 1,000 swimmers estimated by the USEPA for the geometric-mean *E. coli* standard for bathing water of 126 col/100 mL compares well with illness rates predicted from USGS and ORSANCO data. Therefore, on the basis of Ohio data, the geometric-mean bathing-water standard for *E. coli* provides a somewhat similar level of protection regardless of predictive equation used.

Table 7.--Percentage of U.S. Geological Survey and Ohio River Valley Water Sanitation Commission samples that exceeded *Escherichia coli* water-quality standards while meeting fecal-coliform standards

[Effective from May 1 through October 15. All values are in colonies per 100 milliliters. USGS, U.S. Geological Survey; ORSANCO, Ohio River Valley Water Sanitation Commission.]

Standard	Pooled USGS data	Pooled ORSANCO data
Geometric mean:[a]		
Bathing waters[b]	23.9	0.0
Primary contact[c]	58.5	25.4
Single sample:[d]		
Bathing waters[b]	16.1	0.0
Primary contact[c]	39.0	13.3
Secondary contact[e]	31.9	6.6

[a] The geometric mean is based on a minimum of five samples in a 30-day period.

[b] Bathing waters are suitable for swimming and other full-body-contact exposure where a lifeguard or bathhouse is present.

[c] Primary-contact waters are suitable for full-body contact such as swimming, canoeing, and scuba diving.

[d] This value cannot be exceeded in more than 10 percent of the samples collected in a 30-day period.

[e] Secondary-contact waters are suitable for partial-body contact, such as wading.

The *E. coli* geometric-mean criteria for bathing waters and primary-contact recreation are the same (126 col/100 mL); however, the fecal-coliform standards increase fivefold from bathing water (200 col/100 mL) to primary-contact recreation (1,000 col/100 mL). If predicted *E. coli* concentrations are substituted into equation 1, the level of risk rises substantially from bathing waters to primary-contact recreation. For the pooled USGS data, a predicted *E. coli* concentration of 790 col/100 mL nearly doubles the illness rate from 9.4 illnesses per 1,000 swimmers for bathing waters to 15.5 illnesses per 1,000 swimmers for primary-contact recreation. For pooled ORSANCO data, a predicted *E. coli* concentration of 484 col/100 mL raises the illness rate from 7.1 illnesses per 1,000 swimmers (bathing waters) to 13.5 illnesses per 1,000 swimmers (primary-contact recreation). Therefore, the geometric-mean bathing-water and primary-contact *E. coli* criteria do not provide a similar level of protection on the basis of the level of risk.

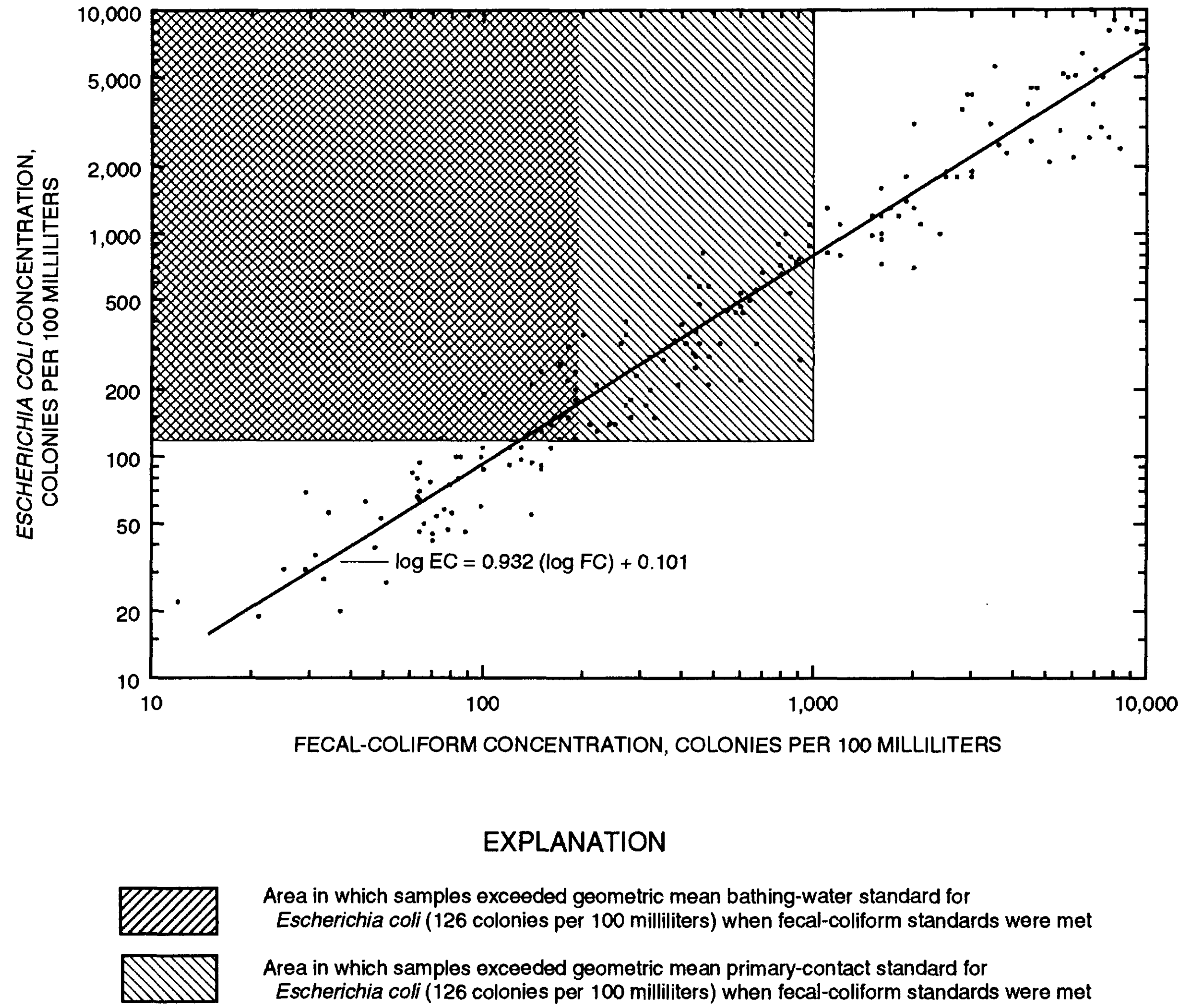
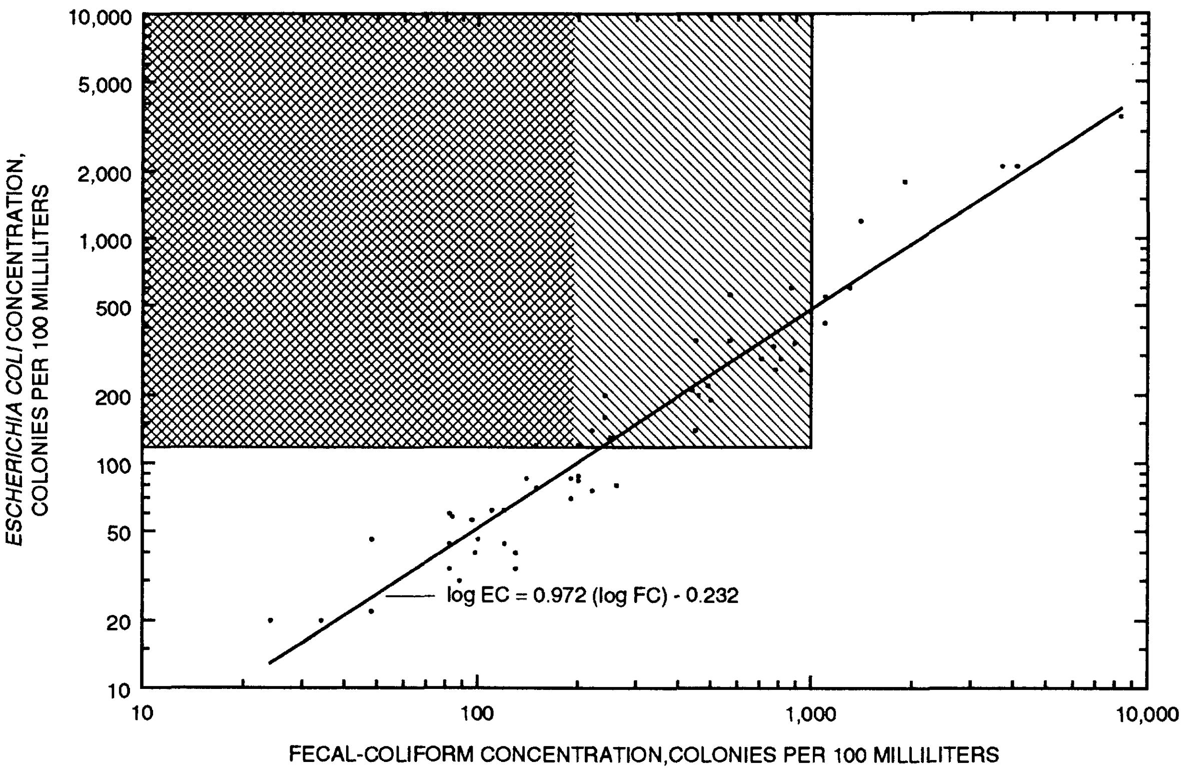


Figure 7.--Samples that exceeded Ohio water-quality standards for *Escherichia coli* when fecal-coliform standards were met for bathing waters and primary-contact recreation: evaluated by use of U.S. Geological Survey data and geometric-mean standards.



EXPLANATION

Area in which samples exceeded geometric mean bathing-water standard for *Escherichia coli* (126 colonies per 100 milliliters) when fecal-coliform standards were met

Area in which samples exceeded geometric mean primary-contact standard for *Escherichia coli* (126 colonies per 100 milliliters) when fecal-coliform standards were met

Figure 8.--Samples that exceeded Ohio water-quality standards for *Escherichia coli* when fecal-coliform standards were met for bathing waters and primary-contact recreation: evaluated by use of Ohio River Valley Water Sanitation Commission data and geometric-mean standards.

SUMMARY AND CONCLUSIONS

In 1986, the USEPA recommended that *E. coli* be used in place of fecal-coliform bacteria as the indicator of fecal pollution in State recreational water-quality standards. This recommendation was based on the results of two studies in which investigators found a statistically significant relation between the rate of swimming-associated gastrointestinal illness and *E. coli* concentration but not fecal-coliform concentration. Results of other epidemiological studies reviewed for this study show that *E. coli* is superior to fecal-coliform bacteria as a predictor of swimming-associated gastroenteritis.

The state of Ohio has temporarily adopted both indicators into State recreational water-quality standards; however, water-resource managers in Ohio have decided to examine fecal-indicator data from Ohio waters and review other pertinent literature to decide whether to use *E. coli* or fecal-coliform bacteria as the basis for State recreational water-quality standards. Data collected in Ohio by the USGS and ORSANCO were statistically analyzed to determine if the presence of *E. coli* could be predicted consistently from fecal-coliform concentrations. If so, then measurement of both may be unnecessary.

Analysis of USGS and ORSANCO data and review of other water-quality studies showed *E. coli* and fecal-coliform concentrations to be highly correlated. In addition, results of regression analysis of USGS and ORSANCO data showed that *E. coli* concentrations could be predicted fairly accurately from fecal-coliform concentrations in Ohio waters. Because *E. coli* concentrations are related to gastroenteritis rate and can be predicted from fecal-coliform concentrations, it would seem, therefore, that fecal-coliform concentration should also be related to gastroenteritis rate. Because researchers found this was not always the case, further statistical analysis of the EC/FC relation was done.

A strong correlation between *E. coli* and fecal-coliform bacteria in natural waters does not ensure that the relation between these two indicators is the same for all waters. For example, results from other studies and analysis of USGS and ORSANCO data from Ohio streams indicate that regression equations derived for prediction of *E. coli* concentrations from fecal-coliform concentrations differ among data sets representing different areas and collected by different agencies. In addition, researchers found different EC/FC ratios ranging from 0.283 to 1.59. In Ohio, predicted EC/FC ratios were intermediate to those found in the literature and ranged from 0.46 to 0.51 for ORSANCO data and 0.72 to 0.89 for USGS data.

ANCOVA on regression equations generated from USGS and ORSANCO data showed that the data could be pooled into two data sets, one for USGS data and one for ORSANCO data; however, ANCOVA also showed that the regression equations generated from data from the two agencies were significantly different. As a result, USGS and ORSANCO data could not be combined for statistical analysis.

Several factors could contribute to the variations observed in regression relations and EC/FC ratios among data sets collected from different locations by different agencies. One factor could be the growth and enumeration of non-fecal coliforms on fecal-coliform plates at some sites where thermotolerant non-fecal coliforms were present in high concentrations. Indeed, there are site differences between ORSANCO and USGS sites; USGS data were collected from inland rivers in Ohio, whereas ORSANCO data were collected from the Ohio River. Another factor could be a dominant contaminant source at some sites or some other environmental factor that strongly influences the EC/FC ratio.

To further test the consistency of the EC/FC relation, the investigators determined 95-percent confidence intervals and an upper 90-percent prediction value for regression relations generated by USGS and ORSANCO data. When fecal-coliform concentrations were set to equal geometric-mean bathing-water and primary-contact water-quality standards for Ohio, predicted *E. coli* geometric-mean values and 95-percent confidence intervals were above the corresponding *E. coli* standard for both recreational-use designations by use of USGS data but below the *E. coli* standard for bathing waters but not for primary-contact waters by use of ORSANCO data. Ninety-percent prediction intervals also showed that *E. coli* predictions calculated from the USGS-generated EC/FC regression equations were different and resulted in different water-quality assessments than predictions based on ORSANCO-generated EC/FC regression equations.

In addition to differences in the regression relation between *E. coli* and fecal coliform, analysis of USGS and ORSANCO data indicated that the percentage of samples that exceeded Ohio *E. coli* standards while attaining fecal-coliform standards differed considerably between agencies. Generally, attainment of Ohio water-quality standards was more difficult when *E. coli* concentrations were used to assess recreational quality than when fecal-coliform concentrations were used. Other investigators have suggested that the new USEPA *E. coli* criteria are more difficult to attain than current State fecal-coliform water-quality standards. The risks of illness calculated by substituting predicted *E. coli* concentrations into the USEPA regression equation, however, do not seem to be substantially different using USGS and ORSANCO data, if comparisons are made within use designations.

Variations in sampling or laboratory techniques between USGS and ORSANCO workers could affect reported *E. coli* and fecal-coliform concentrations; however, the purpose of this study was not to determine actual fecal-indicator concentrations but rather to examine EC/FC relations at different sites. Because both the USGS and ORSANCO used standard methods, the EC/FC relation should remain consistent if fecal-coliform bacteria were to be used instead of *E. coli* in State recreational water-quality standards. Based on the results in this report, the EC/FC relation does not provide the necessary degree of consistency on a statewide basis that is needed to justify the use of fecal-coliform criteria in place of *E. coli* criteria in State standards.

The difference between the use of *E. coli* and fecal-coliform bacteria is that *E. coli* can be used to establish guidelines and standards on the basis of an acceptable level of risk as determined by a regulatory agency and the public. The relation between fecal-coliform bacteria and *E. coli* concentrations can vary, whereas the epidemiological literature shows that the relation between *E. coli* and swimming-associated illness is strong and consistent over geographic boundaries.

REFERENCES CITED

American Public Health Association, American Water Works Association, and Water Pollution Control Federation, 1989, Standard methods for the examination of water and wastewater (17th ed.): Washington, D.C., American Public Health Association, [variously paged].

Bordner, R.H., Winter, J.A., and Scarpino, P., eds., 1978, Microbiological methods for monitoring the environment, water and wastes: Cincinnati, Ohio, U.S. Environmental Protection Agency, EPA-600/8-78-017, 33 p.

Britton, L.J., and Greeson, P.E., eds., 1989, Methods for collection and analysis of aquatic biological and microbiological samples: U.S. Geological Survey Techniques of Water-Resources Investigations, book 5, chap. A4, pt. 1, p. 3-41.

Cabelli, V.J., 1981, Health effects criteria for marine recreational waters: Cincinnati, Ohio, U.S. Environmental Protection Agency, EPA-600/1-80-031.

Campbell, L.M., Michaels, G., Klein, R.D., and Roth, I.L., 1976, Isolation of *Klebsiella pneumoniae* from lake water: Canadian Journal of Microbiology, v. 22, p. 1762.

Caplenas, N.R., Kanarek, M.S., and Dufour, A.P., 1981, Source and extent of *Klebsiella pneumoniae* in the paper industry: Applied and Environmental Microbiology, v. 42, p. 779.

Dufour, A.P., 1984, Health effects criteria for fresh recreational waters: Cincinnati, Ohio, U.S. Environmental Protection Agency, EPA-600/1-84-004, 33 p.

Dufour, A.P., and Cabelli, V.J., 1976, Characteristics of *Klebsiella* from textile finishing plant effluents: Journal of Water Pollution Control Federation, v. 48, p. 872.

Fandrei, G.L., 1985, 1984 Mississippi River bacteria study: Minneapolis, Minn., Minnesota Pollution Control Agency, Division of Water Quality, Monitoring and Analysis Section, 29 p.

Fattal, B., Peleg-Olevsky, E., Agursky, T., and Shuval, H.I., 1987, The association between seawater pollution as measured by bacterial indicators and morbidity among bathers at Mediterranean bathing beaches of Israel: Chemosphere, v. 16, nos. 2 and 3, p. 565-570.

Ferley, J.P., Zmirou, D., Balducci, F., Baleux, B., Fera, P., Larbaigt, G., Jaco, E., Moissonnier, B., Blineau, A., and Boudot, J., 1989, Epidemiological significance of microbiological pollution criteria for river recreational waters: International Journal of Epidemiology, v. 18, no. 1, p. 198-205.

Gannon, J.J., and Busse, M.K., 1989, *E. coli* and enterococci levels in urban stormwater, river water and chlorinated treatment plant effluent: Water Research, v. 23, no. 9, p. 1167-1176.

Holmes, P.R., 1989, Research into health risks at bathing beaches in Hong Kong: Journal of the Institution of Water and Environmental Management, v. 3, no. 5, p. 488-495.

Iman, R.L., and Conover, W.J., 1983, A modern approach to statistics: New York, John Wiley, 497 p.

Looney, S.W., and Gulledge, T.R., 1985, Use of the correlation coefficient with normal probability plots: The American Statistician, v. 39, p. 75-79.

Lorenz, R.C., Hsu, J.C., and Tuovinen, O.H., 1982, Performance variability, ranking, and selection analysis of membrane filters for enumerating coliform bacteria in river water: Journal of the American Water Works Association, August 1982, p. 429-437.

Milligan, J.D., 1987, Assessment of *E. coli*, enterococci, and fecal-coliform bacteria in a recreational floatway: Lake and Reservoir Management, v. 3, p. 163-171.

Ohio Environmental Protection Agency, 1990, Numerical and narrative criteria for recreational use designations: Ohio Administrative Code, Chapter 3745-1, May 1, 1990, p. 07-07 and 07-34.

Seyfried, P.L., Tobin, R.S., Brown, N.E., and Ness, P.F., 1985, A prospective study of swimming-related illness, II—morbidity and the microbiological quality of water: American Journal of Public Health, v. 75, no. 9, p. 1071-1075.

Sladek, K.J., Suslavich, R.V., Sohn, B.I., and Dawson, F.W., 1975, Optimum membrane structures for growth of fecal coliform organisms, *in* EPA/ASTM Symposium on Recovery of Indicator Organisms Employing Membrane Filters, Fort Lauderdale, Florida, January 20-21, 1975, Reprint, 6 p.

Stevenson, A.H., 1953, Studies of bathing water quality and health: American Journal of Public Health, v. 43, p. 529-538.

U.S. Environmental Protection Agency, 1985, Test method for *Escherichia coli* and enterococci in water by the membrane-filter procedure: Cincinnati, Ohio, U.S. Environmental Protection Agency, EPA/600/4-85-076, 24 p.

———1986a, Bacteriological ambient water-quality criteria; availability: Federal Register, v. 51, no. 45, p. 8012-8016.

———1986b, Ambient water-quality criteria for bacteria—1986: Washington, D.C., U.S. Environmental Protection Agency, EPA-A440/5-84-002, 18 p.

Ward, J.R., and Harr, C.A., 1990, Methods for collection and processing of surface-water and bed-material samples for physical and chemical analyses: U.S. Geological Survey Open-File Report 90-140, p. 8.