Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/watres

Association of gastrointestinal illness and recreational water exposure at an inland U.S. beach

J.W. Marion [a], J. Lee [a,b], S. Lemeshow [a], T.J. Buckley [a,*]

[a] Division of Environmental Health Sciences, College of Public Health, The Ohio State University, 320 W. 10th Ave., A-333C Starling-Loving Hall, Columbus, OH 43210, United States

[b] Department of Food Science & Technology, The Ohio State University, Columbus, OH 43210, United States

ARTICLE INFO

Article history:
Received 2 April 2010
Received in revised form
25 June 2010
Accepted 20 July 2010
Available online 29 July 2010

Keywords:
Inland beach
Swimming-associated illness
Water quality
Exposure
Faecal indicator

ABSTRACT

Recent epidemiology studies examining U.S. recreational water exposure and illness relationships have focused primarily on coastal and Great Lakes beaches. Human-made lakes in the U.S. have received little attention in epidemiology studies despite contributing to more waterborne disease epidemics annually than coastal U.S. waters. In a comprehensive beach cohort study, we examined relationships between water quality indicators and reported adverse health outcomes among users of a beach at an inland U.S. lake. Human health data was collected over 26 swimming days during the 2009 swimming season in conjunction with water quality measurements. Adverse health outcomes were reported 8–9 days post-exposure via a phone survey. Wading, playing or swimming in the water was observed to be a significant risk factor for GI illness (adjusted odds ratio (AOR) of 3.2; CI 1.1, 9.0). Among water users, *Escherichia coli* density was significantly associated with elevated GI illness risk where the highest *E. coli* quartile was associated with an AOR of 7.0 (CI 1.5, 32). GI illness associations are consistent with previous freshwater epidemiology studies. Our findings are unique in that our observations of positive associations with GI illness risk are based upon a single daily *E. coli* measurement. Lastly, this study focused on an understudied issue, illness risk at inland reservoirs. Our results support the usefulness of *E. coli* as a health-relevant indicator of water quality for this inland U.S. beach.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The link between adverse health outcomes and recreational swimming and bathing has been well documented since the 1950s (Stevenson, 1953). During the last 30 years, associations between faecal indicators and illness have been established for beach swimmers (Prüss, 1998; Wade et al., 2003, 2006; Colford et al., 2007) including the 1984 U.S. EPA study (Dufour, 1984) that established *Escherichia coli* as a criteria indicator for faecal contamination in freshwater (U.S. EPA, 1986).

Furthermore, the current U.S. recreational water quality criteria for faecal indicators are scheduled for revision no later than October 15, 2012 per a consent decree stemming from Natural Resources Defense Council v. Stephen L. Johnson and U.S. EPA (2008). At the present time, there is limited epidemiologic data for recreational water exposure from inland U.S. waters with most of the research in the last twenty-six years having focused on the Great Lakes (Wade et al., 2006; Wade et al., 2008) and coastal marine waters (Wade et al., 2003). Additionally, there has been limited published epidemiologic research evaluating illness outcomes among users of inland recreational waters.

Beyond epidemiologic studies, outbreak analysis provides a basis for evaluating the public health threat and disease

* *Corresponding author*. Tel.: +1 614 293 7161; fax: +1 614 293 7710.
E-mail address: tbuckley@cph.osu.edu (T.J. Buckley).

0043-1354/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.watres.2010.07.065

burden. The evidence from reported outbreaks suggests the majority of recreational waterborne diseases from untreated water outbreaks are from untreated inland waters. In the most recent U.S. Centers for Disease Control (CDC) surveillance report on recreational waters, inland waters contributed to 19 of the 20 (95%) U.S. recreational waterborne disease outbreaks from non-*Vibrio* agents in untreated waters from 2005 to 2006 (Yoder et al., 2008). This finding has been observed in previous CDC surveillance reports. For example in 2003–2004, all 19 of 19 outbreaks (100%) were from inland waters (Dziuban et al., 2006). In 2001–2002, all 11 of 11 outbreaks from untreated waters were non-coastal (Yoder et al., 2004). This evidence is clear in indicating that U.S. non-coastal waters are a public health concern warranting focused epidemiologic investigation.

Prior to 2003, the three most notable freshwater epidemiology studies in the U.S. were Stevenson (1953), Dufour (1984) and Calderon et al. (1991) as cited in Wade et al. (2003), with 5124, 21,777 and 144 participants, respectively. These studies report associations between gastrointestinal (GI) illness and water quality as indicated by faecal indicator bacteria; however, these studies are few and were limited to few bodies of water.

In response to a growing number of arguments against the use of enterococci and *E. coli* standards in the scientific literature, Wade et al. (2003) engaged in a comprehensive meta-analysis of the 27 most relevant epidemiologic studies, which captured studies ranging from 247 to 26,686 participants. For GI illness, the correlation coefficient was $r = 0.86$ for the natural log relative risk of GI illness as a function of *E. coli* density. The Wade et al. (2003) study provided evidence in support of the U.S. EPA's position on using *E. coli* as a recreational water standard to reduce GI illness risk among freshwater contact recreation users. Despite good correlation between *E. coli* density and GI illness risk, the correlation was based solely upon three studies.

To contribute knowledge to the understudied issue regarding illness risk among inland recreational water users, we conducted a prospective cohort study. The primary goal of this study was to evaluate the effectiveness of *E. coli* as an indicator of GI illness risk among recreational water users at East Fork Lake, Ohio, U.S.A. East Fork Lake, is a typical Midwestern U.S. flood-control reservoir used often for recreational activities. This study was designed to identify differences in GI illness risks among swimmers and non-swimmers. By sampling over 26 days, the study design allowed the exploration of GI illness risk for swimmers exposed to various *E. coli* densities.

2. Materials and methods

2.1. Overview of methods

Beach water samples and human health data were collected from the public beach at East Fork State Park during the 2009 swimming season. The study design was a prospective cohort design similar to that used by Wade et al. (2006) and Wade et al. (2008). In brief, subjects were recruited from the beach on the same day that water quality measurements including *E. coli* were conducted. Subjects completed a survey at the beach giving demographics and water-related behaviours (e.g. wading, playing or swimming in water). Subjects were contacted eight to nine days later by phone to ascertain potential water-related illness. The study design, questionnaires and related materials were reviewed and approved by the Institutional Review Board at The Ohio State University (IRB Protocol #2009H0107).

2.2. Sampling site

Beach water sampling and the human health survey were administered at the public beach at East Fork State Park (39°1′ 11.2″ N; 84° 1′ 2.8″ W) over 26 weekend days starting May 30 and ending August 30, 2009. The 365 m public beach is located along the lake, a human-made 8.7 km^2 flood-control reservoir providing numerous benefits to the Cincinnati, Ohio, U.S.A., metropolitan/suburban area including full-body contact water recreation. Probable pollution sources for the reservoir are primarily non-point, with dominant watershed land uses being row crop agriculture (37%), light urban/residential (33%) and forestland (25%) (East Fork Watershed Collaborative, 2006). Several small municipal wastewater treatment plants are permitted to discharge into the lake tributaries; however, no discharges are permitted directly into the reservoir. Failing home sewage treatment systems have been identified as likely faecal contamination sources for the lake tributaries and reservoir (East Fork Watershed Collaborative, 2006).

2.3. Recreational water samples

Samples were collected in accordance with the Ohio Department of Health (2009) standard methods. In brief, a single daily sample was collected using autoclaved 500 mL Nalgene bottles or 500 mL sterile Whirl-Pak© bags in water of approximately three feet in depth near the center of the 365 m beach. The sample was gathered by sweeping the bottle or bag one foot below the surface of the water. Scum and debris were avoided in the sweeping process. From this single daily sample collection at the beach, all of the analyses and archiving were performed. Within-day sample times varied from 10:50 to 20:25 although most samples (92%) were collected within 4 h of the median sample time of 13:52. Ancillary water measurements and weather observations were made at the time of sampling. Daily precipitation and the mean daily water inflow for the watershed of the reservoir were provided for each sample collection day by the Louisville District of the U.S. Army Corps of Engineers (2010). All water quality laboratory analyses were performed within 6 h of sample collection. *E. coli* density was quantified using EPA Method 1603 (U.S. EPA, 2002). Autoclaved deionised water was used as a negative control. The colony counts were converted to colony forming units (CFU)/100 mL after dividing by the filtration volume.

2.4. Administration of beach user questionnaire

The sampling approach was a cluster design that enrolled households with a household member over 18 years old who was using the East Fork beach and willing to speak on behalf of their household. Subjects were recruited over 26 days across

13 weekends during the summer of 2009. Beachgoers were invited into the study through signage and/or by being approached by study personnel. The invitation for participation followed an approved standard script. The process of consent was administered verbally to persons eligible to consent. Upon receiving consent, households choosing to engage in the full study participated in an initial interview at the beach, a second interview when leaving the beach, and follow-up phone interview 8–9 days after their visit.

The initial interview at the beach captured pertinent enrolment information. This included the identification of household members at the beach as well as phone numbers for the follow-up telephone interview. In addition, information was collected concerning potentially confounding existing symptoms as well as age and sex. This initial interview was completed with a gift and instructions for reporting for second interview to occur before leaving the beach.

The second interview occurred at the end of their beach stay and gathered household member-specific exposure data including: time spent at the beach, time in the water, whether the body or head were submerged, consumption of food at the beach, and others. This second interview concluded with a gift and information concerning the telephone interview follow-up survey.

The telephone interview was the third and final interview. Using previously collected telephone numbers, the participants were contacted 8–9 days following their visit to the East Fork beach. Respondents were asked about any new symptoms since their visit to the beach.

The survey instrument employed was a modified version of the U.S. EPA questionnaire developed for the National Epidemiological and Environmental Assessment of Recreational Water (NEEAR) study. The full questionnaire has been successfully employed by Wade et al. (2006), Wade et al. (2008) and Heaney et al. (2009). Like the original NEEAR Beach Questionnaire, our modified and reduced length questionnaire retained key questions to ascertain three types of information: exposure status, illness status/symptoms, and demographics. Across all three information domains, the identified adult respondent ($\geq$18 years of age) provided information for themselves and others within their household. Information about exposure status included the amount of time spent at the beach, as well as their activities including the amount of time spent in the water, head immersion, and others. Illness/symptom questions included a variety of outcomes including GI symptoms, ear infection, skin rash or skin infection. Demographic questions were limited to age, gender and ethnicity. Ancillary information was collected including distance travelled to visit the beach, frequency of visits, and whether cost (i.e. no charge) was a factor in their decision to visit the beach.

2.5. *Classification of beach exposures*

To characterize beach exposures, dichotomous and categorical classifications were used. The dichotomous classifications pertained to the identification of exposed (coded as "1") versus unexposed ("0"). The exposed swimmer group represented respondents that reported to "wade, swim or play in the water"; whereas, the unexposed group indicated no such activity. An additional exposure of interest was the consumption of food at the beach. Respondents reporting to have "consume[d] food while at the beach" the day of the survey were also dichotomously classified as "1" or "0".

A categorical classification scheme was used to further characterize the level of microbial exposure among the swimmer group. The water quality parameter *E. coli* was used for class coding. For the logistic regression model, the *E. coli* classes were 0–3.3, $\geq$3.3–11.3, $\geq$11.3–59 and $\geq$59–1538 CFU/100 mL. These values are the quartiles for the *E. coli* distribution of the beach water samples among the 806 swimmers.

2.6. *Classification of health outcomes*

The data were explored to identify any reported illness, GI illness and highly credible gastrointestinal illness (HCGI). Symptoms including but not limited to fever, headache, nausea, diarrhoea, and vomiting were treated as dichotomous variables, where individuals reporting a symptom were coded as "1" for the symptom or "0" if no symptom was reported. GI was defined using the definition established by Dufour (1984), which includes any person reporting any of the following symptoms: nausea, stomachache, diarrhoea or vomiting. GI illness was treated as a dichotomous variable, where individuals with no new GI illness symptoms were coded as "0" and persons with one or more new GI illness symptoms since visiting the beach were coded as "1". HCGI was defined using the "HCGI-1" definition employed by Colford et al. (2007) and nearly identical to the HCGI illness definition established by Dufour (1984). Persons reporting to exhibit any of the following conditions were coded as "1" for HCGI illness: (1) vomiting; (2) diarrhoea and fever; (3) stomachache or nausea accompanied with a fever.

2.7. *Data analysis*

GI illness was initially explored by performing tabulations with exposure and demographic data. Initial data analysis focused on the occurrence of specific GI-related symptoms among swimmers and non-swimmers. Further tabulations were performed to compare the frequency of GI and HCGI illnesses among swimmers and non-swimmers. GI and HCGI illness frequencies were further explored across the four levels of water quality as determined by the distribution of the *E. coli* density quartiles. Additional tabulations for characterizing the frequency of GI-related symptoms among beach users were performed taking into consideration age, sex and other beach exposures such as food consumption.

Following initial data exploration, multivariable analysis was employed. Logistic regression was used to estimate the risk of GI and HCGI illness via observed adjusted odds ratios. Models were constructed to consider potential confounders and/or modifying influences such as age and sex. Due to the clustering of participants by household, the data were not considered to be a simple random sample, but instead for the purposes of statistical analysis, the data were treated as clustered by household (Levy and Lemeshow, 1999). Model construction involved the identification of biologically plausible covariates coupled with statistical selection. A reverse stepwise procedure was used in model 3 to select variables with some likely association ($p < 0.15$; Bendel and Afifi (1977)) with GI and HCGI illness. Model 3 estimates adjusted odds ratios for swimmers across the four *E. coli* density levels.

Ultimately, three adjusted logistic regression models were constructed. The first model (model 1) estimated GI illness risk for swimmers, where swimmers are those who reported wading, playing or swimming in the water. This model used non-swimmers as the reference group and adjusted for only age, gender and the reservoir inflow. With concerns that age and flow were not likely to be linear in the logit for GI and HCGI illness outcomes, age was categorized into six groups that could be classified as young child, older child, teenager, young adult, adult and older adult. Reservoir inflow was categorized using the terciles of exposure for the respondents. The consumption of food at the beach and varying levels of *E. coli* were not included in this model, despite being associated with the illness outcomes. The second model (model 2) estimates GI and HCGI illness risk for persons who consumed food at the beach. Model 2 was constructed similar to model 1, with adjustments being made for only age, gender and reservoir inflow. Model 2 did not differentiate swimmers from non-swimmers and did not account for *E. coli* density, albeit associated with GI illness. The third model (model 3) is the most complex. Model 3 only included swimmers and was developed to determine GI and HCGI illness risk among swimmers in waters with various densities of *E. coli*. Model 3 only included age, gender and the statistically-associated covariates ($p < 0.15$). The dichotomous variable regarding food consumption at the beach and the categorical variable for the 48 h reservoir inflow were both included due to their level of association with GI and HCGI illness in this model. For all models, an adjusted odds ratio (AOR) not including 1.0 in the 95% CI of the AOR was considered significant, where $p < 0.05$. The performance of each of the models was evaluated based on fit (Hosmer-Lemeshow Goodness-of-Fit Test (Hosmer and Lemeshow, 2000)) and an analysis of model sensitivity/specificity. Model sensitivity/specificity was assessed by interpreting the area under the receiver-operating-characteristic (ROC) curve (Metz, 1978).

Linear regression procedures were also employed to identify the strength of association between GI illness and *E. coli* density. Linear regression and Pearson correlation analysis was performed as in Dufour (1984) to permit comparison with previous epidemiologic data. Using this methodology, *E. coli* (CFU/100 mL) was plotted on a log scale versus GI and HCGI illness per 1000 individuals. Three dates containing less than 10 total observations from swimmers were excluded due to their small sample size and large standardized residual. Additionally, three of the remaining 23 sampling dates had mean *E. coli* densities less than 1 CFU/100 mL and were converted to 1 CFU/100 mL. Pearson correlations were determined to be significant if $p < 0.05$.

Associations between illness rates (GI and HCGI) and *E. coli* density were also evaluated for trend. The Cochran–Armitage Test was employed where GI and HCGI illness rates were on an ordinal scale from lowest mean daily *E. coli* density to highest. Trends were determined using the Cochran–Armitage Test for Trend where a $p < 0.05$ indicated a statistically significant trend in the observed data.

Initial tabulations, logistic regression modelling, assessing logistic model fit, and the Cochran–Armitage Test for Trend were performed with Stata 11 (Stata Corporation, College Station, TX, USA). Linear regression modelling and Pearson correlation analysis were performed using Minitab 15 (Minitab Inc., State College, PA, USA).

3. Results

3.1. Enrolment

A total of 682 households consented and completed the initial beach survey. Of these households, 554 returned to complete the beach exit interview. Of those completing the beach exit survey, 300 households were successfully contacted and completed the follow-up phone survey for an overall household study retention rate of 44%. Accordingly, results are presented for the 300 households and 965 individuals who were living at the household for the entire follow-up period. Demographic characteristics of this group are provided in Table 1. Our sample was predominantly white (93%), and travelled less than 40 miles to visit the beach (81%).

3.2. Characterization of beach water quality

A total of 26 water samples were collected over 13 weekends and all were successfully analyzed. The beach was open and occupied over the 26-day sampling period. For two of the 26 sample days, *E. coli* densities exceeded the advisory threshold of 235 CFU/100 mL. For these days, densities of 1538 CFU/100 mL and 487 CFU/100 mL were detected, representing the highest observed values (Table 2). These two high densities were observed during the same weekend and followed 6.7 cm of rainfall the day before the first observation. *E. coli* densities remained above advisory conditions for a second day even

Table 1 – Individual and household characteristics of persons surveyed at East Fork Lake (Ohio, United States) with complete surveys and telephone follow-up.

Household/respondent characteristics	No. (%)
Household size (No. of individuals)	
1	51 (17)
2	66 (22)
3	55 (18)
4	64 (21)
5	37 (12)
≥6	27 (9.0)
Distance travelled to beach (Miles)	
0–9.9	170 (18)
10–19.9	242 (25)
20–29.9	192 (20)
30–39.9	177 (18)
≥40	184 (19)
Ethnicity of individual household members	
White	895 (93)
Black	22 (2.2)
Hispanic	16 (1.7)
Asian	9 (0.9)
American Indian	5 (0.5)
Other	1 (0.1)
Missing	17 (1.8)

Table 2 – Descriptive statistics of beach water quality and beach usage during sampling days ($N = 26$) at East Fork Lake (Ohio, United States).

Beach parameter	N	Mean ± S.E.	1st Quartile	Median	3rd Quartile	Range
Temperature (°C)	26	26.7 ± 0.4	25.5	25.9	28.6	24.2–30.4
Specific conductivity (μS/cm)	26	276 ± 2.3	268	273	282	259–307
Turbidity (NTUs)	26	22.6 ± 4.7	10.6	14.6	21.3	7.0–116
E. coli (CFU/100 mL)	26	95.1 ± 60.7	0.5	9.1	37.8	0–1538
48 h reservoir inflow (m^3/sec)	26	504 ± 313	0	52	351	0–8184
Beach user density (Persons)	26	185.0 ± 27.3	57.8	156	302	4.0–483.0

with no additional rainfall. Tributary inflows of 214 m^3/s and 18.2 m^3/s respectively, were observed on these two sample days. The largest 24 h inflow volume reported during the entire study period was 214 m^3/s. Median inflow for this reservoir during the sampling period was 1.5 m^3/s. Beach water *E. coli* densities were generally low, as illustrated by a median value of 9.1 CFU/100 mL.

3.3. *Characterizing exposures of beach users*

Characteristics of individuals completing the survey with respect to both demographics and exposure activities are provided in Table 3. These data indicate that males and females were similar in their behaviours to wade, play or swim in the water (80 vs. 77%), but females were more likely to consume food at the beach (64 vs. 57%). The youngest age group (0–5 years) was most likely to consume food at the beach with 64% of group members having this exposure. Additionally, 90% of the youngest age group members had beach water exposure. The older children (6–11 years) and the adolescent/teenager group (12–18 years) had the most reported exposure to the water; whereby 98% of 6–11 year olds and 93% of 12–18 year olds reported to wade, play or swim. The senior age group (56.0–73.9 years) reported the least frequency for food consumption (38%), and wading, playing or swimming in the water (41%).

Table 3 – Summary of beach exposures and reported illness by gender and age classification of respondents during the 2009 swimming season at East Fork Lake (Ohio, United States).

Beach users attributes	No. (%)	Beach exposures		Reported illnesses	
		No. consumed food at beach (%)	No. exposed body to water (%)	GI illness	HCGI illness
Gender					
Female	540 (56)	333 (64)	444 (82)	23 (4)	10 (2)
Male	425 (44)	241 (57)	362 (85)	25 (6)	3 (1)
Age (years)					
0–5	127 (13)	81 (64)	114 (90)	11 (9)	2 (2)
6–11	174 (18)	102 (59)	170 (98)	8 (5)	2 (1)
12–18	137 (14)	85 (62)	128 (93)	5 (4)	3 (2)
19–30	174 (18)	103 (59)	147 (84)	8 (5)	1 (1)
31–55	306 (32)	182 (59)	224 (73)	15 (5)	5 (2)
56–74	32 (3)	12 (38)	13 (41)	1 (3)	0 (0)
Missing	15 (2)	9 (60)	10 (67)	0 (0)	0 (0)

3.4. *Characterizing illness among beach users*

Of the 965 individuals included in our sample, there were 109 cases (11.3%) of reported total illness. Of these, 48 (44%) were GI-related. The partitioning of these adverse health outcomes across demographic and exposure factors are presented in Tables 3–5. Table 3 illustrates that GI illness incidence was similar among males and females; however, the youngest age group had the highest incidence of GI illness. Among the exposure activities, Table 4 demonstrates the group not exposed to water reported 3 (1.9%) cases. Only swimmers in the lowest *E. coli* density exposure group reported less GI illness (1.1%), although no statistical difference exists between these groups. GI illness proportions were significantly higher for swimmers in the two highest *E. coli* density groups, and in these two groups over 8% of swimmers reported GI illness. Stomachaches, diarrhoea, and HCGI illness were also reported more frequently in these highest two groups.

3.5. *GI illness risk estimates for beach users*

Using logistic regression, we developed three models to evaluate predictors for GI and HCGI illness risk among East Fork beach users. Table 5 provides the adjusted odds ratios (AOR) for three models. The results of model 1, adjusted for age, sex and reservoir inflow, demonstrated a significant AOR for persons who reported wading, swimming or playing in the water (AOR = 3.2; CI: 1.1, 9.0) suggesting a 3.2-fold increased odds for GI illness among those who went in the water. The AOR for HCGI was not determined as the non-swimmer reference group reported no HCGI cases. Model 2, adjusted for age, sex and reservoir inflow, demonstrated increased odds for GI illness (AOR = 3.6; CI: 1.4, 9.9) and HCGI illness (AOR = 7.2; CI: 1.1–48) for persons reporting to have consumed food while at the beach.

Model 3, which only estimates GI and HCGI illness risk among persons who had contact with the water, demonstrated significant GI illness risk estimates for persons exposed to the highest *E. coli* levels. After adjusting for age, sex, reservoir inflow, and consuming food at the beach, exposure to beach waters in the second highest *E. coli* quartile (>11.3–59 CFU/100 mL) presented a significant elevated odds ratio for GI illness (AOR = 7.2; CI: 1.3, 39), but not for HCGI illness (AOR = 6.0; CI: 0.54, 71). Likewise, the highest *E. coli* quartile (>59–1551 CFU/100 mL) presented similar findings for GI (AOR = 7.0; CI: 1.5, 32) and HCGI illness (AOR = 3.7; CI: 0.63–77). Although statistical significance was not achieved in the odds ratios for HCGI illness in these two highest *E. coli*

Table 4 – Summary of gastrointestinal illnesses for beach users across various *E. coli* density exposure levels at East Fork Lake (Ohio, United States) during the 2009 swimming season.

	Not exposed to water	*E. coli* density levels among exposed individuals			
		0–3.3 CFU/100 mL	≥3.3–11.3 CFU/100 mL	≥11.3–59 CFU/100 mL	≥59–1538 CFU/100 mL
No. of individuals	159	186	178	208	234
GI symptom	No. w/symptom (%)	No. w/symptom (%)	No. w/symptom (%)	No. w/symptom (%)	No. w/symptom (%)
Nausea	0 (0)	0 (0)	3 (1.7)	2 (0.96)	6 (2.6)
Vomit	0 (0)	1 (0.54)	1 (0.56)	3 (1.4)	6 (2.6)
Diarrohea	3 (1.9)	2 (1.1)	4 (2.2)	16 (7.7)	12 (5.1)
Stomachache	0 (0)	0 (0)	0 (2.8)	1 (0.48)	8 (3.4)
Fever	0 (0)	0 (0)	5 (2.8)	3 (1.4)	4 (1.7)
GI illness	3 (1.9)	2 (1.1)	6 (3.4)	18 (8.7)	19 (8.1)
HCGI illness	0 (0)	1 (0.54)	1 (0.56)	5 (2.4)	6 (2.6)

density classes, the respective *p*-values ($p = 0.152$ and $p = 0.287$) demonstrate that the direction of an association, if it exists, is towards increased odds of HCGI illness. Similarly, for the lowest *E. coli* density class (>3.3–11.3 CFU/100 mL) above the reference group (0.1–3.3 CFU/100 mL), the adjusted odds ratio did not achieve statistical significance ($p = 0.204$); however, the likely direction of the association, if one exists, is towards increased odds of GI illness.

3.6. Assessing performance of models and trends

The three logistic regression models constructed to predict GI and HCGI illness vary in performance with respect to sensitivity and specificity as indicated by the ROC curve. Using the area under the ROC curve (AUC), the value for the GI illness model in model 1 is 0.63 (Table 6). Since the AUC value is below 0.70, we conclude the model does not provide acceptable discrimination (Hosmer and Lemeshow, 2000). The models used to estimate GI and HCGI risks in model 2 provide better discrimination as indicated by AUC values of 0.69 and 0.76, respectively. The AUC value of 0.76 provides acceptable discrimination. The most discriminatory models are the GI and HCGI illness risk models labelled as model 3 in Table 6. The AUC value of 0.72 in the GI illness risk model provides fair or acceptable discrimination; however, the AUC value of 0.81 in the HCGI illness risk model provides good discrimination. Additionally, when assessing the fit of each model, all of the *p*-values are greater than 0.05 when using the Hosmer-Lemeshow Goodness-of-Fit Test. These *p*-values from the goodness-of-fit test suggest that the probabilities estimated by the model are an accurate representation of the true disease experienced in the data.

Significant trends were observed in the data. The Cochran–Armitage Test for Trend demonstrated a significant ordinal trend among swimmers for GI illness ($p = 0.0016$) and HCGI illness ($p = 0.0151$) with increasing *E. coli* densities. Furthermore, the continuous *E. coli* data plotted on a logscale demonstrates a positive association with increasing

Table 5 – Association of beach user attributes with illness outcomes across 26 sampling days at East Fork Lake (Ohio, United States).

Model	No. (%)	GI			HCGI		
		No. Ill (% Ill)	AOR (95% CI)	*p*-value	No. Ill (% Ill)	AOR (95% CI)	*p*-value
Model 1: Exposed body to water[a]							
Unexposed	159 (16)	3 (1.9)			0 (0)		
Exposed[b]	806 (84)	45 (5.6)	*3.2 (1.1–9.0)*	*0.028*	13 (1.6)	–[c]	
Model 2: Consumed food at beach[a]					45 (5.6)		
Unexposed	379 (40)	8 (2.1)			7 (1.9)		
Exposed	574 (60)	40 (7.0)	*3.6 (1.4–9.9)*	*0.010*	36 (6.3)	*7.2 (1.1–48)*	*0.040*
Model 3: *E. coli* density (CFU/100 mL)[d]							
0.1–3.3	186 (23)	2 (1.1)			1 (0.5)		
>3.3–11.3	178 (22)	6 (3.4)	3.2 (0.53–19)	0.204	1 (0.6)	1.6 (0.07–37)	0.768
>11.3–59	208 (26)	18 (8.7)	*7.2 (1.3–39)*	*0.022*	5 (2.4)	6.0 (0.54–71)	0.152
>59–1551	234 (29)	19 (8.1)	*7.0 (1.5–32)*	*0.013*	6 (2.6)	3.7 (0.63–77)	0.287

a Adjusted for age, sex, reservoir inflow, and clustering within households.
b Exposure is defined as persons who reported to wade, swim or play in the water.
c Unable to report AOR for HCGI in exposed group since no cases of HCGI were reported in the unexposed group.
d Adjusted for age, sex, reservoir inflow and consuming food at beach, and clustering within households after excluding persons reporting no wading, swimming or playing in water.

Table 6 – Model performance values for three logistic regression models predicting illness (GI and HCGI illness) across 26 swimming season days at East Fork Lake (Ohio, United States).

Model	GI		HCGI	
	Area under ROC curve	Hosmer-Lemeshow Goodness-of-Fit Test (p-value)	Area under ROC curve	Hosmer-Lemeshow Goodness-of-Fit Test (p-value)
Model 1: Exposed body to water[a,b]	0.63	0.3605	–[c]	–[c]
Model 2: Consumed food at beach[a]	0.69	0.1772	0.76	0.8454
Model 3: *E. coli* density (CFU/100 mL)[d]	0.72	0.0714	0.81	0.7788

a Adjusted for age, sex, reservoir inflow, and clustering within households.
b Exposure is defined as persons who reported to wade, swim or play in the water.
c Unable to report since no cases of HCGI were reported in the unexposed group.
d Adjusted for age, sex, reservoir inflow and consuming food at beach, and clustering within households after excluding persons reporting no wading, swimming or playing in water.

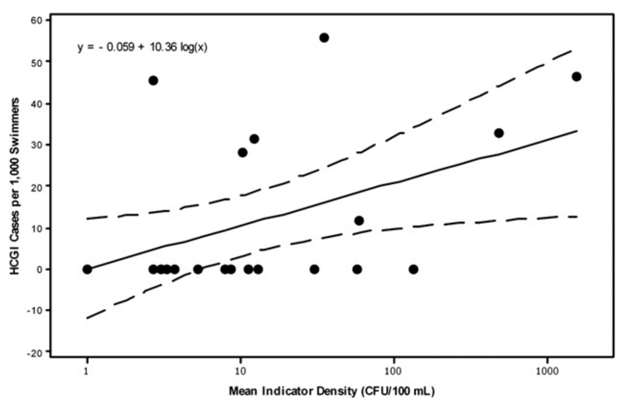

Fig. 1 – Regression line estimating highly credible gastrointestinal illness rate among swimmers at East Fork Lake (Ohio, United States) exposed to varying *E. coli* densities across 23 days[1].

proportions of GI and HCGI cases among swimmers. Fig. 1 illustrates the positive association between HCGI cases and increasing *E. coli* density, which is supported by the correlation coefficient ($r = 0.467$) and the significant correlation p-value ($p = 0.025$) listed on Table 7. Table 7 further describes the significant correlation between GI illness frequency and *E. coli* density where $p = 0.031$.

4. Discussion

The study showed increased GI illness risk among swimmers when compared to non-swimmers, consistent with inland (Stevenson, 1953; Dufour, 1984; Wade et al., 2006) and marine beach studies (Cabelli et al., 1979; Colford et al., 2007). Furthermore, this study demonstrated the effectiveness of *E. coli* as a faecal indicator for determining GI illness risk among swimmers at the study beach. The association of GI illness and the results of the single daily faecal indicator measurement is an important finding that compliments the results of Wade et al. (2006), in which a single rapid *Enterococcus* measurement collected in the morning was useful for determining GI illness risk among swimmers.

This prospective cohort study demonstrated an association between contact recreation with beach water and illness. Overall, swimmers were more likely to report experiencing symptoms than non-swimmers, a finding consistent with Dufour (1984) and Fleisher et al. (1996). Like Dufour (1984), GI symptoms were the most frequently reported. The observed GI illness incidence of 56 cases per 1000 among swimmers is comparable to the inland reservoir beach study results from Dufour (1984) where the range of GI illness incidence over the two-year study period was 37.9–61 cases per 1000 swimmers. Non-swimmers at the inland reservoir studied by Dufour (1984) exhibited a range in GI illness of 19–53 cases per 1000; whereas, East Fork non-swimmers experienced 19 GI cases per 1000. Similarly, the proportions of swimmers reporting HCGI illness were comparable. Overall, the rates of GI illness in this study are comparable to previously reported results from the largest epidemiologic study of inland U.S. beaches.

The increased risk of GI illness was also observed to be associated with persons reporting to have consumed food at the beach. This variable was identified through reverse selection when constructing the logistic regression model and presumed to be a variable that serves as a proxy for duration of exposure at the beach while taking into account confounding food-related illnesses.

The logistic regression models constructed demonstrate significant associations with GI and HCGI illness (Table 5); however, only after taking into account the *E. coli* densities does the model achieve acceptable model sensitivity and specificity (AUC > 0.70). Table 6 shows that the best performing model is model 3 when the outcome variable is HCGI (AUC = 0.81). This result, in which the HCGI illness model outperformed the GI illness outcome model, is consistent with the results of Dufour (1984). The most plausible reason

[1] Sample days with less than 10 observations were removed from the analysis due to their large standardized residual, resulting in a total of three of the 26 days being removed from the regression analysis.

Table 7 – Summary statistics of linear regression between GI illness outcomes and *E. coli* density across 23 sampling days[a] at East Fork Lake (Ohio, United States).

Symptom	Slope	Y-intercept	Std. error of slope	Correlation coefficient	Correlation *P*
GI	27.51	12.57	13.08	0.451	0.031
HCGI	10.36	−0.06	4.28	0.467	0.025

a Sample days with less than 10 observations were removed from the analysis due to their large standardized residual, resulting in a total of three of the 26 days being removed from the regression analysis.

is the unmistakeable nature of HCGI illness versus more subjective GI symptoms, which enables improved classification of the truly ill. A all three models provided an accurate representation of the reported disease as demonstrated by the Hosmer-Lemeshow Goodness-of-Fit Test (Table 6). Model 3, however, has a *p*-value near 0.05 for GI outcome suggesting less than an ideal fit and may indicate that an alternative model may produce a better representation of the true disease. In contrast, the more well-defined HCGI outcome for model 3, suggests a very accurate representation of the reported disease.

The observations of increasing GI and HCGI illness incidence in the higher *E. coli* density exposure groups when compared with swimmers exposed to the least *E. coli* suggest that with increasing *E. coli* density, the risk for GI and HCGI illness increases. These results are confirmed by significant AORs when GI illness is the outcome of interest. The HCGI illness outcome was unable to achieve significance, likely as an artefact of the small sample size and relatively constant and favourable water quality. Our models are all limited in that the sample size is small and potentially confounding variables, such as additional recreational water exposures, were not able to be considered due to the abbreviated questionnaire administered. Despite these limitations, using swimmers exposed to the lowest *E. coli* level as a reference group, the assumption is that exposures away from the beach would be similar across all swimmers, regardless of the day they selected to swim at the study lake. Additionally, linear regression analysis of swimmers demonstrated a significant positive correlation between GI and HCGI incidence with increasing *E. coli* density (Table 7). This strength of association was further upheld for both GI and HCGI incidence in trend tests.

Given the limited resources, we were only able to collect a single water quality sample per day. The time of sample collection was not consistent throughout the day, but was still able to produce a significant association with GI and HCGI illness. It has been determined that tremendous variability can exist in water quality at selected marine beaches (Boehm et al., 2002); however, in this study, we were able to estimate GI illness risk using a single daily sample. These results may not be able to be generalised to reservoirs with different watershed dynamics or pollution sources; however, this study lake had stable reservoir inflows and precipitation during the days in which participants were enrolled. Additional *E. coli* measurements are not available to challenge our assumption that *E. coli* densities were stable throughout our individual study days.

Lastly, this study reaffirmed the problem of not having a rapid tool available to prevent exposure to recreational waters when infectious disease risk is great enough to warrant an advisory, as recreational water users were unknowingly exposed to advisory conditions on two occasions. The sample day with the peak reservoir inflow was associated with beach user exposures to waters with *E. coli* densities exceeding 1500 CFU/100 mL. Despite having limited data regarding water quality and reservoir inflows at this reservoir, it appears that predictive modelling of *E. coli* at this reservoir offers an opportunity to inform recreational water users in advance of advisory events in an approach similar to Francy et al. (2003). Predictive modelling efforts of *E. coli* are underway at this reservoir. Future studies using archived filter membranes from the study period are being planned to enable quantification of alternative indicators using rapid molecular techniques.

5. Conclusions

- Beach users who reported wading, swimming or playing in the water at this inland beach were at greater risk for GI illness than persons who did not report these behaviours.
- The risk of GI illness increased among swimmers exposed to increasing densities of the faecal indicator bacteria, *E. coli*.
- A single *E. coli* sample per day was effective at estimating infectious disease risk at our study beach and may be an acceptable approach at similar inland lake beaches.
- The membrane filtration method for *E. coli* quantification continues to be inadequate in providing beach users with timely information regarding infectious disease risk, and rapid or predictive tools are warranted.

Acknowledgements

Funding for this research was provided by a grant from the Ohio Water Development Authority. Jason Marion received fellowship support through the Public Health Preparedness for Infectious Diseases targeted investment in excellence at The Ohio State University. We are appreciative of the staff with Ohio State Parks, especially Scott Fletcher, for supporting the study. We are grateful to the East Fork Lake beachgoers who graciously agreed to participate in our survey and to fellow graduate (Michael Lofton, James Rosenblum, Pei-Yu Chiang, Kaedra Wetzel) and undergraduate (Sara Zolinski and Matthew Farley) students who gave up their multiple summer weekends to assist in recruiting subjects, completing surveys, and entering data.

REFERENCES

Bendel, R.B., Afifi, A.A., 1977. Comparison of stopping rules in forward stepwise regression. Journal of the American Statistical Association 72 (357), 46–53.

Boehm, A.B., Grant, S.S., Kim, J.H., Mowbray, S.L., McGee, C.D., Clark, C.D., Foley, D.M., Wellman, D.E., 2002. Decadal and short period variability of surf zone water quality at Huntington Beach, California. Environmental Science and Technology 36 (18), 3885–3892.

Cabelli, V.J., Dufour, A.P., Levin, M.A., McCabe, L.J., Haberman, P.W., 1979. Relationship of microbial indicators to health effects at marine bathing beaches. American Journal of Public Health 69 (7), 690–696.

Calderon, R., Mood, E., 1991. Health effects of swimmers and nonpoint sources of contaminated water. International Journal of Environmental Health Research 1, 21–31.

Colford, J.M., Wade, T.J., Schiff, K.C., Wright, C.C., Griffith, J.F., Sandhu, S.K., Burns, S., Sobsey, M., Lovelace, G., Weisberg, S., 2007. Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. Epidemiology 18 (1), 27–35.

Dufour A., 1984. Health Effects Criteria for Fresh Recreational Waters. U.S. Environmental Protection Agency Report EPA-600-1-84-2004. Cincinnati, OH.

Dziuban, E.J., Liang, J.L., Craun, G.F., Hill, V., Yu, P.A., Painter, J., Moore, M.R., Calderon, R.L., Roy, R.L., Beach, M.J., 2006. Surveillance for waterborne-disease outbreaks associated with recreational water – United States, 2003–2004. Mortality and Morbidity Weekly Report 55 (12), 1–24.

East Fork Watershed Collaborative, 2006. East Fork Lake Tributaries Watershed Management Plan Online Document. URL: ftp://ftp.co.clermont.oh.us/oeq/EastForkLakeTributariesWAP_draft_9_25_06.pdf.

Fleisher, J.M., Day, D., Salmon, R.L., Jones, F., Wyer, M.D., Godfree, A.F., 1996. Marine waters contaminated with domestic sewage: nonenteric illnesses associated with bather exposure in the United Kingdom. American Journal of Public Health 86, 1228–1234.

Francy, D.S., Gifford, A.M., Darner, R.A., 2003. *Escherichia coli* at Ohio Bathing Beaches – Distribution, Sources, Wastewater Indicators, and Predictive Modeling. U.S. Geological Survey, Water-Resources Investigations Report 02-4285, Columbus, OH.

Heaney, C.D., Sams, E., Wing, S., Marshall, S., Brenner, K., Dufour, A.P., Wade, T.J., 2009. Contact with beach sand among beachgoers and risk of illness. American Journal of Epidemiology 170 (2), 164–172.

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression, second ed. Wiley, New York.

Levy, P.S., Lemeshow, S., 1999. Sampling Populations: Methods and Applications, third ed. Wiley, New York.

Metz, C.E., 1978. Basic principles of ROC analysis. Seminars in Nuclear Medicine 8, 283–298.

Natural Resources Defense Council v. Stephen L. Johnson, U.S. EPA, 2008. United States District Court for the Central District of California No. CV06-4843 PSG (JTLx). Filed 04 September 2008.

Ohio Department of Health, 2009. Beach Monitoring Online document. URL: http://www.odh.ohio.gov/odhprograms/eh/bbeach/beachmon.aspx.

Prüss, A., 1998. Review of epidemiological studies on health effects from exposure to recreational water. International Journal of Epidemiology 27 (1), 1–9.

Stevenson, A.H., 1953. Studies of bathing water and health. American Journal of Public Health 43, 529–538.

U.S. Army Corps of Engineers, Louisville District, 2010. Yearly Lake Reports Online database. URL: http://www.lrl.usace.army.mil/wcds/dlbrpt.htm.

U.S. Environmental Protection Agency, 1986. Ambient Water Quality Criteria for Bacteria U.S. Environmental Protection Agency Report EPA-440/5-84-002. Washington DC.

U.S. Environmental Protection Agency, 2002. Method 1603: *Escherichia coli* (*E. coli*) in Water by Membrane Filtration Using Modified Membrane-Thermotolerant *Escherichia coli* Agar (Modified mTEC) U.S. Environmental Protection Agency Report EPA-821-R-02-023, Washington DC.

Wade, T.J., Pai, N., Eisenberg, J.N., Colford, J.M.J., 2003. Do U.S. Environmental protection agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. Environmental Health Perspectives 111, 1102–1109.

Wade, T.J., Calderon, R.L., Sams, E., Beach, M., Brenner, K.P., Williams, A.H., Dufour, A.P., 2006. Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestinal illness. Environmental Health Perspectives 114 (1), 24–28.

Wade, T.J., Calderon, R.L., Brenner, K.P., Sams, E., Beach, M., Haugland, R., Wymer, L., Dufour, A.P., 2008. High sensitivity of children to swimming-associated gastrointestinal illness: results using a rapid assay of recreational water quality. Epidemiology 19 (3), 375–383.

Yoder, J.S., Blackburn, B.G., Craun, G.F., Hill, V., Levy, D.A., Chen, N., Lee, S.H., Calderon, R.L., Beach, M.J., 2004. Surveillance for waterborne-disease outbreaks associated with recreational water – United States, 2001–2002. Mortality and Morbidity Weekly Report 53 (8), 1–22.

Yoder, J.S., Hlavsa, M.C., Craun, G.F., Hill, V., Roberts, V., Yu, P.A., Hicks, L.A., Alexander, N.T., Calderon, R.L., Roy, S.L., Beach, M.J., 2008. Surveillance for waterborne diseases and outbreaks associated with recreational water use and other aquatic facility-associated health events – United States, 2005–2006. Mortality and Morbidity Weekly Report 57 (9), 1–29.