International Journal of Hygiene and Environmental Health 277 (2026) 114888

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

A Randomised Controlled Trial Assessing Infectious Disease Risks from Bathing in Inland Recreational Waters

Edward K.S. Lam [a,*], Márta Vargha [b], Daniel Thomas [c], Roland Salmon [a,1], Mihály Kádár [a,2], Maria José Figueras Salvat [d], Julii Brainard [a], Jamie Bartram [e,f], Zsofia Barna [a,3], Paul R. Hunter [a]

[a] Norwich Medical School, University of East Anglia, Norwich, NR4 7TJ, UK
[b] National Center for Public Health and Pharmacy, Budapest, 1097, Hungary
[c] Communicable Disease Surveillance Centre, Public Health Wales, Cardiff, CF10 4BZ, UK
[d] University Rovira I Virgili, Avenida Catalunya s/n, Tarragona, 43002, Spain
[e] School of Civil Engineering, University of Leeds, Leeds, LS2 9BT, UK
[f] Department of Environmental Science and Engineering, School of Public Health, University of North Carolina at Chapel Hill, USA

ARTICLE INFO

Keywords:
Open water swimming
Rivers
Lakes
Escherichia coli
Faecal enterococci
Somatic coliphages
Gastrointestinal illness

ABSTRACT

Evidence to determine the suitability of water quality standards to prevent illness from recreation exposure in inland waters is limited. We present a secondary analysis of previously unpublished data from four Hungarian freshwater sites included in the Epibathe study, a large randomised controlled trial. A total of 2368 participants were randomly allocated either to bathe for 10 min undertaking at least three head immersions, or to remain on the shoreside without water contact. Concurrent water sampling quantified individual-level faecal indicator organism densities and health outcomes were assessed at one-week follow-up.

Generalised estimating equation models quantified the relative risk for adverse health outcomes for bathers versus shoreline-only (non-bather) participants, as well as the change in risk per unit increase in FIO concentration; crude, covariate-adjusted and covariate-adjusted models with multiple imputation are reported. Higher concentrations of *Escherichia coli* and somatic coliphage were each associated with increased risk of gastrointestinal illness (adj. RR = 1.73; 95% CI 1.13–2.65 and RR = 1.46; 95% CI 1.04–2.04 respectively). Bathing itself, independent of any individual microbial indicator, was associated with increased risk for skin ailments (adj. RR = 2.17; 95% CI 1.55-3.03). Low prevalence of eye, ear or respiratory infections precluded reliable estimation of exposure-response relationships.

These findings confirm the value of *E. coli* and potential of somatic coliphage densities as indicators of freshwater quality relevant to recreation-associated gastrointestinal illness risk. In freshwater settings, *E. coli* and coliphages appear to be more informative than enterococci as predictors of gastrointestinal illness, contrasting with evidence from marine waters.

1. Introduction

Outdoor recreation on inland-waterways is a popular pastime (University of Brighton, Exegesis SDM Ltd, Rubicon Associates, Plumpton College and G. a. L. Hughes, 2015; RYA, 2023). Although engagement with natural "blue spaces" has been associated with diverse health and economic benefits (Gascon et al., 2017; White et al., 2020; Cromley and Mackintosh, 2021; Geneshka et al., 2021), direct contact with, and in particular ingestion of, untreated freshwater may also pose a risk for waterborne infections. Microbial contamination remains a persistent problem for many freshwater bathing sites (Binns, 2023) and recreational exposure has been linked to disease outbreaks. For example, a single open water river swimming event on the River Thames, UK, in October 2012 was associated with 338 cases of gastrointestinal illness

* Corresponding author.
E-mail address: edward.lam@uea.ac.uk (E.K.S. Lam).
[1] Consultant Epidemiologist (retired).
[2] Retired; previously National Institute of Environmental Health, Hungary.
[3] Independent consultant on water hygiene, https://legionellakontroll.hu/.

https://doi.org/10.1016/j.ijheh.2026.114888
Received 5 June 2026; Received in revised form 3 August 2026; Accepted 16 August 2026
Available online 22 August 2026

linked in part to *Cryptosporidium* and *Giardia* infections (Hall et al., 2017).

Human and animal faecal material may enter freshwater systems from a range of sources, including sewage discharges, agricultural and urban surface runoffs, wildlife and human activities, including bathing itself. Assessment of recreational water quality typically relies on the quantification of faecal indicator organisms (FIOs), which act as proxy measures for faecal contamination and the potential presence of enteric pathogens. There is a strong evidence base for FIO-based monitoring in marine contexts, where randomised controlled trials have shown robust associations between concentrations of faecal streptococci and gastrointestinal illness (Cabelli et al., 1982; Dufour, 1984; Ferley et al., 1989; Kay et al., 1994).

Although FIO-based monitoring is widely used by regulatory frameworks, there is no consensus on which indicator organisms are most appropriate for freshwater environments. The European Bathing Water Directive (2006/7/EC; European Union, 2006) retains both *E. coli* and intestinal enterococci standards, with separate threshold values for inland and coastal waters. In contrast, the most recent iteration of the World Health Organisation's Guidelines on Recreational Water Quality (WHO, 2021) has removed *E. coli*, solely retaining intestinal enterococci for both marine and freshwaters. This divergence reflects ongoing uncertainty regarding how well FIO concentrations predict health risks in freshwater environments dominated by diffuse, non-point contamination (Kozak et al., 2025). Observational cohort studies, comparing illness outcomes among self-selected bathers and non-bathers, have attempted to address this uncertainty. Whilst informative, these studies are vulnerable to selection and confounding bias, especially when baseline behaviours and susceptibility differ between groups and are not adequately controlled for (Kay et al., 1994). Additionally, individual-level exposure is often poorly characterised, with limited use of concurrent microbiological measurements. As a result, cohort studies have reported inconsistent associations between exposure to FIO concentrations and illness risk (Fewtrell and Kay, 2015).

Addressing this evidence gap requires studies that pair individual-level FIO exposure measurements with health outcomes under randomised controlled conditions; to date, however, only one such randomised controlled trial (RCT) has been published (Wiedenmann et al., 2005).

Here we report a secondary analysis of previously unpublished individual-level data from the four freshwater sites included in the Epibathe study, a large randomised controlled study carried out in 2006/2007 at bathing sites in Europe. Participants were prospectively recruited and randomly allocated to bathing or non-bathing exposure groups. All bathers had comparable bathing duration and immersion behaviours. Concurrent water sampling enabled the characterisation of individual-level FIO exposure and health outcomes were assessed at one and three-week follow-up. Following Wiedenmann et al. (2005), we restricted subsequent analyses to outcomes reported at one-week follow-up, consistent with the acute time course expected of water-associated illnesses. Using population-average generalised estimating equation models that adjusted for socio-demographic covariates and accounted for clustering by study site, we quantified associations between specific faecal indicator concentrations and subsequent illness risk. Drawing upon data from four study sites spanning a broad microbial exposure gradient, these findings strengthen the epidemiological basis for *E. coli* and somatic coliphages as indicator organisms for recreational water-quality guidelines in freshwater environments affected by diffuse contamination.

2. Methods

A detailed experimental 'manual' for undertaking the Epibathe study was developed and revised as the trials proceeded. The manual contained instructions and guidance for trialists and support staff, including lessons learned, steps taken to ensure data quality, reasons for aspects of the protocol and other protocol information. This manual is available in Supplementary Text 1 (Supplementary Materials). A shorter summary of trial implementation is below.

The study protocol (ethical clearance no. 230-70/2006-1018EKU) was approved by the institutional ethics committee of the Fodor József National Public Health Centre and the Hungarian Scientific and Research Ethics Committee prior to participant recruitment. All participants gave written informed consent prior to any involvement in the study.

2.1. Study sites

The Hungary field studies were a mixture of two riverine and two lacustrine locations. They were chosen in order to capture the natural variability in water movement, mixing, usage and potential for microbial contamination across bathing sites in Central Europe. All four sites were compliant with the *Imperative* standard of the EU Bathing Water Directive (76/160/EEC). Under this Directive, compliant sites required 95% of seasonal samples not exceed the imperative values of 2000 faecal coliforms and 10,000 total coliforms per 100 mL, with guide values of 100 faecal coliforms (80% of samples) and 100 faecal streptococci per 100 mL (90% of samples). The trial sites are described individually and specifically in the Methods (next paragraphs), but their respective results are only indicated by non-identifiable numbers otherwise (numbered from 10 to 13). Further details and maps are available in Supplementary Text 3.

Dömsöd and Fadd bathing sites were predominantly lacustrine in nature, situated on fossil reaches of the River Danube: dead-arm channels that are hydrologically connected to, but largely isolated from, the main river. The Fadd site was more hydrologically enclosed and exhibited no direct water exchange with the main Danube channel, but there was limited exchange at the Dömsöd site, particularly during flooding. Neither site received direct effluents from wastewater treatment facilities, but both sites were subjected to diffuse sources of contamination associated with local human settlements, recreational huts and waterborne activity, including bathing and fishing.

Tiszakécske and Csongrád riverine sites were situated along the Tisza River, a transboundary river that flows through Romania and Ukraine before entering Hungary. In contrast with the partially isolated fossil reaches in Dömsöd and Fadd, these riverine bathing areas are directly connected with upstream waters, with continuous freshwater flow and less opportunities for stagnation. The Tisza River from the Tiszakécske site receives secondary treated sewage effluent both directly and indirectly through small water inflows (Fózer et al., 2024). Both human and animal (bovine and porcine) faecal input have been identified (Rusiñol et al., 2014). At the Csongrád site, water quality at the bathing site is primarily influenced by diffuse pollution from agricultural activity, holiday homes, unofficial bathing sites, bathing and fishing.

Meteorological and water conditions differed appreciably between trial days. In the 2006 season, the Dömsöd trial was conducted on 16/07/2006 in sunny weather with air and water temperatures both of 25–26°C, whereas the Fadd trial on 13/08/2006 was overcast and cooler, with air temperature in the low 20 s °C and water at 18–19°C. In 2007, the Csongrád trial took place on 01/07/2007 in hot weather (air above 30°C, water 25–26°C) and the Tiszakécske trial on 05/08/2007 on a cool, cloudy day (air approximately 20°C, water 18–19°C).

2.2. Microbiological testing

The four study sites were divided at 10-m intervals into six bathing zones. Water samples were collected at 20-min intervals throughout the 3-h trial period, giving nine sampling rounds and 54 samples per trial session at each site. Bathing commenced 10 min before the first sampling round and ended 10 min after the last. Samples were taken simultaneously from all six zones, in the middle of each zone at the point of exposure, 30 cm below the surface at chest depth (140 cm, or 90 cm in

the children's zone), into sterile 1-L flasks. In accordance with ISO19458:2006, samples were transported to the laboratory on melting ice within 2 h of collection, stored at 4°C on arrival and processed between three and 8 h after sampling.

These samples were analysed by the laboratory of the Hungarian National Public Health Center for *E. coli*, Intestinal Enterococci and Somatic coliphage: levels were quantified using the microtitre 96 well Most Probable method (EC96) (ISO9308-3:1998), microtitre 96 well Most Probable method (IE96) (ISO7899-1:1998) and plaque assay (ISO10705-2:2000) respectively.

The limits of detection were 3.5 MPN/100 mL for both *E. coli* (EC96) and intestinal enterococci (IE96), analysed at 1:2 dilutions in triplicate, and 7 PFU/100 mL for somatic coliphages, from 5 mL plated in triplicate. To avoid undefined values at zero, non-detects were assigned a value of half the smallest observed non-zero titre prior to log-transformation.

2.3. Recruitment and randomisation of participants

During initial conceptualisation of the Epibathe study in 2006, no formal power calculations were performed: minimum sample sizes were determined with reference to GI illness alone with target sizes set by reference to prior freshwater (Wiedenmann et al., 2005) and marine (Kay et al., 1994) studies, in which statistically significant associations between *E. coli* titres and GI illness risk was detected with 829:920 and 548:668 bathers to non-bathers respectively. Overall, 1133 bathers and 1235 non-bathers were recruited for Epibathe, which exceeded this design benchmark.

Volunteers were predominantly recruited from the local communities situated within 15-50 km of each study site. Recruitment was facilitated by internet and radio advertisements, poster campaigns and through collaboration with local public health offices. Standardised baseline questionnaires collected from volunteers included pertinent information on demographics, recent medical history and potential behavioural confounders. All answer forms were scanned into a participant dataset using optical character reading (OCR). Steps taken to ensure data quality are documented in the trial manual in Supplementary Text 1. Volunteers were allowed to participate if they gave informed consent, appeared to be generally in good health, were willing to bathe at the sites and said that they were able to comply with study procedures. Each volunteer who gave complete information (filled in all questionnaires) was eligible to receive 40 € compensation in shopping vouchers.

On the trial day, volunteers were again interviewed to account for any changes in recent medical history or potential behavioural confounders since recruitment. Using a computer-generated allocation schedule in Microsoft Excel, volunteers were randomly allocated to one of two groups: bathers or non-bathers. Bathers and non-bathers were separated and identified distinctively by specific colour wrist bands or t-shirts. All bathers were required to bathe once for 10 min within their designated zone, making at least three full facial immersions. Individual exposure levels (of faecal indicators) were assigned based on their zone and nearest time zone where bathing took place; the mean value was used when exposure occurred exactly between two time zones. Non-bather volunteers were instructed not to bathe one week before and after the trial; during the trial, they were not allowed to bathe at all. In the non-bather zone alternative activities were provided, such as a bouncy castle, football and climbing walls. Non-bathers had to stay on the beach for 3 h while the bathers had their immersions. Only bathers were allowed into the bather area, but every volunteer could go into the non-bather area. Both bathers and non-bathers were supplied with the same sandwich lunch options.

2.4. Health outcomes

Seven days after the trial, participants were interviewed in person or by telephone to assess for post-exposure health conditions and symptoms. Following the approach of Wiedenmann et al. (2005), we defined the five primary health outcomes (below) using Boolean combinations of reported symptoms (operators in capitals). These symptom definitions mirrored those used in Wiedenmann et al. to facilitate comparability.

(i) **Gastrointestinal illness** (GI): *diarrhoea* OR *vomiting* OR (*nausea* AND *fever*) OR (indigestion AND fever)
(ii) **Acute febrile respiratory illness**: (*headache* OR *joint pain* OR *blurred vision* OR *loss of appetite* OR *tiredness* OR *dizziness* OR *pins & needles* OR *muscle cramps*) AND (*sore throat* OR *dry cough* OR *productive cough* OR *shortness of breath* OR *runny nose*)
(iii) Ear infection
(iv) Eye infection
(v) **Skin Ailment**: *skin rash* OR *skin ulcer* OR *itching*

2.5. Predictor variables

For each primary health outcomes, we estimated the independent association with each of the following exposure variables:

(i) Bathing group status
(ii) Microbial water-quality indicators (log-transformed titres):
  a. *E. coli*
  b. Enterococci
  c. Coliphages

These exposures were analysed independently, using separate generalised estimating equation (GEE) models. Additional covariates were considered with each model:

(i) **Prior illness** (three weeks before water exposure): to adjust for pre-existing background prevalence, each model included an indicator of prior illness matching the case definition of the outcome, i.e. prior GI illness when modelling GI illness outcomes.
(ii) **Age group**: categorised as 4-10, 11-20, 21-30, 31-40 and $>$ 40 years.
(iii) **Potential confounders**: we evaluated several behavioural variables that were potentially associated with gastrointestinal infections and other health outcomes, e.g. raw or unpasteurised milk consumption, stomach remedy and alcohol intake (see Tables S1–5 in Supplementary Text 2). Covariates demonstrating evidence of association with the health outcome at the $p < 0.10$ level were considered potential confounders and were retained for inclusion in subsequent multivariate generalised estimating equation (GEE) models. For rare health outcomes, some behavioural variables exhibited perfect separation and were thus excluded from subsequent multivariate GEE models.

2.6. Generalised estimating equations

To estimate the association between bathing-water exposure and adverse health outcomes, we used GEE models to obtain population-average (marginal) relative risks (GEE; Liang and Zeger, 1986). This approach accounts for the clustering of participants within study sites and potential geo-temporal variations in underlying illness prevalence. Unlike random-effects models, GEEs directly estimate marginal effects without requiring assumptions about the distribution of site-level effects or their independence from exposure variables: such assumptions would be violated in environmental studies, where site characteristics are intrinsically linked to microbial water quality.

All analyses were conducted in R version 4.5.1 using the glmtoolbox package (version 0.1.12). All data and code needed to reproduce the analyses are provided in the project Github repository (https://github.com/edwardkslam/Epibathe). Significance threshold was set at $p < 0.05$.

2.6.1. Model structure

Let Y_{ik} denote the binary health outcome, e.g. gastrointestinal illness at one-week follow-up, for participant i in study site (cluster) k:

$$Y_{ik} \sim Bernoulli(p_{ik})$$

With p_{ik} being the probability of illness for participant i in site k.

We modelled this probability using a log-linked binomial GEE, so that

$$\log(p_{ik}) = \beta_0 + \beta_{exp}X_{exp,ik} + \beta_1 X_{1,ik} + \beta_2 X_{2,ik} + \cdots + \beta_p X_{p,ik}$$

where $\beta_{exp}X_{exp,ik}$ and $X_{1,ik}, X_{2,ik}, \cdots, X_{p,ik}$ denotes the exposure and covariate values (see above) respectively for participant i in site k. Under the log-link, exponentiated coefficients yield the relative risks:

$$RR_j = e^{\beta j}$$

Which represents the multiplicative change in risk associated with a one-unit increase in predictor X_j.

Participants within the same site are likely to experience similar environmental factors and baseline exposure to background illness. To account for non-independence of individual responses within each study site, we specified an exchangeable working correlation, meaning that all pairs of participants within the same site share a common correlation parameter α:

$$Corr(Y_{ik}, Y_{i'k}) = \begin{cases} 1 & i = i' \\ \alpha & i \neq i' \end{cases}$$

Robust (sandwich) standard errors (Huber, 1967; White, 1980) were used to ensure valid inference in case the working correlation structure was mis-specified.

2.7. Multiple imputation

To assess for robustness to missing data, we imputed missing covariate values using multiple imputation. Missing data were limited to covariate information (Table S6 in Supplementary Text 2); the fraction of missing information differed across the health outcomes considered, with the highest proportion observed in the GI models (1.65%; n = 39). Following the rule of thumb recommended by White et al. (2011), whereby the number of imputations should exceed the percentage of incomplete cases to ensure reproducible standard errors, we generated $m = 5$ imputed data sets. These were performed under a fully conditional specification model with the Missing At Random (MAR) assumption using the mice package (version 3.18.0) in R. These imputation models included all covariates, exposure variables and health outcomes, in order to preserve underlying associations between predictors and outcomes.

For each imputed data set $d = 1, \cdots, m$, we fitted GEE models and obtained robust estimates for the regression coefficient $\hat{\beta}^{(d)}$ and variance $\hat{V}^{(d)}$. Pooled parameter estimates were derived using Rubin's Rules (Rubin, 1996):

$$\overline{\beta} = \frac{1}{m}\sum_{d=1}^{m}\hat{\beta}^{(d)}\ \mathrm{z}$$

$$\overline{V} = \frac{1}{m}\sum_{d=1}^{m}\hat{V}^{(d)}$$

It is necessary to further inflate within-imputation variance $\overline{V}$ with between-imputation variation B to reflect the effect of missing data and Monte Carlo error.

$$B = \frac{1}{m-1}\sum_{d=1}^{m}\left(\hat{\beta}^{(d)} - \overline{\beta}\right)^2$$

Total (pooled) variance T is therefore given as:

$$T = \overline{V}\left(1 + \frac{1}{m}\right)B$$

Subsequent results from imputed analyses were highly consistent with complete-case estimates (Table 2); given the low proportion of missingness and restriction of imputation to covariates, rather than outcomes, additional sensitivity analyses were not pursued.

3. Results

3.1. Study population

Across the 2006 and 2007 bathing seasons, a total of 2368 persons participated in bathing trials and subsequent one-week follow-up across four freshwater bathing sites in Hungary. Of these, 1133 (47.8%) were randomised to the bather group and 1235 (52.2%) to the non-bathers. Baseline demographic and behavioural characteristics were well-balanced between bathers and non-bathers (Table S7 in Supplemental Text 2): no statistically significant differences were observed across key variables, including prior illnesses. With regards to problem health outcomes, gastrointestinal (n = 48; 2.03%) and skin (n = 47; 1.98%) illnesses had the highest prevalences (Table 1). Outcomes disaggregated by study site (Table S12 in Supplementary Text S2) broadly showed higher incidence rates in bathers across all four sites, though there were site-specific heterogeneities in magnitude.

Initially, 2724 participants were recruited, of whom 2368 completed the one-week follow-up interview, representing an overall retention rate of 86.9% (Table S13 in Supplementary Text S2). Substantial attrition occurred between initial recruitment and the bathing trial (162 across all four sites), most notably at site 13 (645 to 507; −21.4%). This likely reflects the unusually cool conditions on that trial day, which discouraged bathing, rather than any systematic or outcome-related loss. Reassuringly, post-bathing follow-up completion was high across all four sites: sites 10 (679 to 664; −2.2%), 11 (616 to 607; −1.5%), 12 (597 to 592; −0.8%) and 13 (507 to 505; −0.4%). This suggests that outcome ascertainment bias from differential dropout after exposure is unlikely to be a major concern in the analysis.

3.2. Microbiological titres

For bathing participants, we examined study site-specific exposure to the three microbial indicators, which were quantified using standardised laboratory methods (see Methods). Overall, we observed substantial between-site heterogeneity in microbial concentrations (Fig. 1 and Table S8 in Supplementary Text 2).

For sites 10 to 12, *E. coli*, Intestinal Enterococci and somatic coliphage titres were mostly clustered near the lower detection limit with relatively few high-titre observations (Fig. 1 and Table S8 in Supplementary Text 2). Notably, somatic coliphages were not detected at site 11, with all measurements at or below the plaque assay detection limit. In contrast, site 13 exhibited markedly higher and more variable concentrations, particularly for *E. coli* and somatic coliphages. This variability supports the use of site-clustered statistical approaches, in order to assess the associations between microbial indicators and health

Table 1
Frequency of primary health outcomes by bather status.

Health problem	Overall (n = 2368)	Non-bather (n = 1235)	Bather (n = 1133)
Gastrointestinal	48 (2.03%)	24 (1.94%)	24 (2.12%)
Skin	47 (1.98%)	16 (1.30%)	31 (2.74%)
Respiratory	9 (0.38%)	2 (0.16%)	7 (0.62%)
Ear	13 (0.55%)	5 (0.40%)	8 (0.71%)
Eye	13 (0.55%)	4 (0.32%)	9 (0.79%)

Table 2
Relative risks for gastrointestinal illness and skin ailments by exposure variable. Results for respiratory, ear and eye infections are reported separately in Supplementary Text 4.

GI Illness	Crude			Adjusted			Adjusted + MI		
Indicator	RR	95% CI	p-value	RR	95% CI	p-value	RR	95% CI	p-value
Non-bathers	1	-	-	1	-	-	1	-	-
Bathers	1.11	0.61-2.02	0.73	1.13	0.63-2.02	0.68	1.12	0.63-2.00	0.7
Log. *E. coli*	1.48	0.98-2.25	0.06	1.77	1.15-2.73	0.01	1.73	1.13-2.65	0.01
Log. I. Enterococci	1.25	0.67-2.35	0.48	1.20	0.75-1.94	0.45	1.18	0.74-1.88	0.49
Log. S. phage	1.55	1.18-2.04	0.002	1.48	1.05-2.06	0.02	1.45	1.05-2.04	0.03
Skin Ailments	**Crude**			**Adjusted**			**Adjusted + MI**		
Indicator	RR	95% CI	p-value	RR	95% CI	p-value	RR	95% CI	p-value
Non-bathers	1	-	-	1	-	-	1	-	-
Bathers	2.12	1.47-3.06	0.00	2.17	1.55-3.03	0.00	2.30	1.57-3.37	0.00
Log. *E. coli*	0.70	0.36-1.39	0.31	0.61	0.26-1.44	0.26	0.50	0.21-1.16	0.11
Log. I. Enterococci	0.53	0.17-1.58	0.25	0.63	0.24-1.68	0.36	0.50	0.19-1.28	0.15
Log. S. phage	1.38	0.89-2.13	0.15	1.31	0.88-1.93	0.18	1.26	0.79-2.00	0.34

Notes: Relative risks for microbial indicators correspond to a one-unit increase in log-transformed indicator density. RR = Relative risk; Log. = natural logarithmic transformation of raw values of stated indicator; MI = multiple imputation; "Did not converge" means the model could not be constructed; I. Enterococci refers to Intestinal Enterococci; S. phage refers to somatic coliphages.

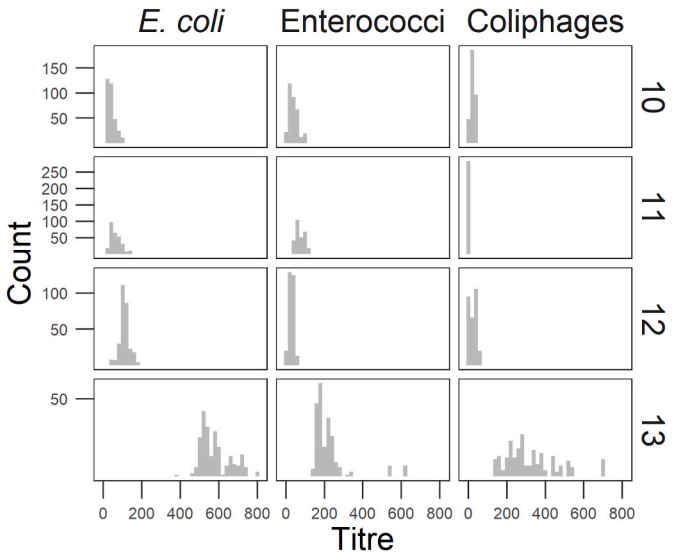

Fig. 1. The distribution of individual-level exposures for bathers across the four Hungarian study sites, as quantified using *E. coli*, Intestinal Enterococci and Somatic coliphage titres. The three indicators were quantified using standardised microbiological methods (see Methods).

outcomes across a broad exposure gradient.

3.3. Gastrointestinal illness

We examined the effect of each of the four exposure variables on gastrointestinal illness in separate GEE models, using only complete cases. After adjusting for recent GI illness, age group and additional behavioural confounders (Table 2, Adjusted Model, middle column), we observed statistically significant positive associations ($p < 0.05$) between GI illness and the microbial indicators *E. coli* (RR = 1.77 per one-unit increase in log-transformed indicator density; 95% CI 1.15-2.73) and somatic coliphage (RR = 1.48; 95% CI 1.05-2.06). In contrast, neither bathing status nor Intestinal Enterococci concentrations showed such statistically significant associations with GI illness.

Notably, the estimated effect of *E. coli* increased after adjustment for age, indicating negative confounding by age, whereby younger participants, particularly 11-20 and 21-30 year olds (Table S9 in Supplementary Text 2), had a higher underlying risk of GI illness than other age groups.

To assess whether the pooled exposure-response estimates were disproportionately influenced by study site 13, which contributed most of the upper exposure range (Fig. 1), we conducted sensitivity analyses by refitting the GI illness models (Table S14 in Supplementary Text 2) with this site excluded. Covariate-adjusted models demonstrated the same positive associations ($p < 0.05$) between GI illness and the indicators *E. coli* (RR = 1.77; 95% CI 1.38–2.28) and somatic coliphage (RR = 1.58; 95% CI 1.14–2.18); the *E. coli* estimate was identical to that from the full dataset and the coliphage estimate marginally larger. The narrower confidence intervals may misleadingly suggest greater precision: indeed, the exclusion of site 13 has removed the most heterogenous cluster. Moreover, with only three remaining study sites, the finite-sample downward bias of the robust sandwich estimator is more pronounced.

We assessed for potential age-specific heterogeneities in illness risk by incorporating interaction terms between log-transformed *E. coli* concentrations and age group in the GEE model (Table S10 in Supplementary Text 2). Although age-specific interaction coefficients were individually significant ($p < 0.05$), the overall interaction term was not (Joint Wald test; $\chi2(3) = 7.41$; $p = 0.06$). Age group-specific average marginal effects (AME) showed modest heterogeneity in the exposure-response relationship (Table S11 in Supplementary Text 2), with larger increases in predicted GI illness risk observed amongst the 11-20 year- (AME = 0.03; 95%CI 0.02-0.03) and 21-30 year- (AME = 0.02; 95%CI –0.01-0.05) old cohorts.

To assess the robustness against missing data, we repeated our analyses with multiple imputation: a total of 39 missing covariate values (1.65% of cases) were imputed, generating 5 replicate datasets, which were subsequently pooled (see Methods). The imputed GEE models (Table 2, Adjusted + MI Model, rightmost column) yielded estimates for the effect size for *E. coli* (RR = 1.73; 95% CI 1.13–2.65) and Somatic coliphage (RR = 1.45; 95% CI 1.05–2.04) that were highly consistent with our initial complete-case analyses, suggesting limited impact from missing data on inference.

3.4. Skin ailments

Skin ailments were the second most common adverse health outcome (n = 47). Co-variate adjusted GEE models showed that bathing status was positively associated with increased risk for skin ailments across both complete-case (RR = 2.17; 95% CI 1.55-3.03) and imputed analyses (RR = 2.30; 95% CI 1.57-3.37; Table 2). In contrast, none of the microbial indicators were associated with increased risk of skin ailments

(at $p < 0.05$).

3.5. Acute febrile respiratory illness, ear and eye infections

Respiratory, ear and eye infections were rare (n = 9, 13 and 13 respectively). The small number of cases resulted in perfect separation for several covariates and prevented stable convergence of the covariate-adjusted models, preventing model convergence and reliable estimation of exposure–response relationships for these outcomes. Full results are reported in Supplementary Text 4 and are not interpreted further here.

4. Discussion

Across all health outcomes considered, GI illness was most consistently associated, as theoretically could be expected, with faecal contamination (Cabelli et al., 1982; Dufour, 1984; World Health Organization, 2021). In particular, higher concentrations of *E. coli* and somatic coliphages were each associated with increased risk (RR = 1.73; 95% CI 1.13–2.65 and RR = 1.45; 95% CI 1.05–2.04 respectively), whereas Intestinal Enterococci showed no such association. Independent of microbial concentrations, bathing itself was associated with an increased risk of skin ailments (RR = 2.30; 95% CI 1.57-3.37). Estimates for the rarer respiratory, ear and eye outcomes were unstable owing to low case numbers and were not interpreted further (Supplementary Text 4).

These results suggest that *E. coli* and somatic coliphages counts are better predictors for GI illness in freshwater environments than enterococci, unlike the case in marine waters (Cabelli et al., 1982; Dufour, 1984; Kay et al., 1994; World Health Organization, 2011). Experimental studies have shown that environmental drivers, such as sunlight, water temperature and interactions with indigenous microbiota, can contribute to differential persistence of FIOs in freshwater and marine waters (Fujioka and Narikawa, 1982; Davies, 1989; Korajkic et al., 2013): notably, *E. coli* and somatic coliphages decay more slowly and persist longer than enterococci in freshwater environments (Anderson et al., 2005), implying that the former may better retain a signal of recent faecal contamination and predict GI health risks. This pattern differs from marine environments, where enterococci exhibit greater environmental stability (Anderson et al., 2005) and have been shown to be reliable predictors of illness, whereas *E. coli* performs poorly as an indicator (Cabelli et al., 1982; Dufour, 1984; Ferley et al., 1989; Kay et al., 1994). Water temperatures across the four trials in this study (18-26°C) fell within the range over which these differences in persistence have been characterised. However, with only a single trial day at each site, the impact of temperature could not be distinguished from other individual- or site-level characteristics.

Taken together, enterococci are likely poor indicators of faecal contamination and associated health outcomes in freshwater settings, emphasising the need for appropriate microbial indicators for inland recreational waters. This has direct implications for current international guidance. In consolidating around intestinal enterococci for both marine and freshwaters, the 2021 WHO Guidelines removed *E. coli* from the recommended indicators for freshwater recreation. Our results do not support such consolidation: intestinal enterococci showed no association with GI illness in freshwaters, whereas *E. coli* showed a consistent exposure-response relationship that persisted under multiple imputation and exclusion of the most contaminated site.

Overall, freshwater and marine systems may require different indicator organisms and compliance regimes to account for the contrasting performance of indicators between environments. The absence of a single unified standard presents regulatory challenges, including the financial and administrative costs of individually classifying bathing locations and assessing them against different standards. Such clear dichotomisation would also fail to provide practical guidance for many popular bathing sites situated within transitional environments, such as estuarine beaches. In these settings, the coexistence of freshwater and marine influences may further complicate the behaviour and predictive value of microbial indicators.

The association between bathing status and skin ailments, in the absence of a corresponding relationship with FIO concentrations, suggests that bathers are exposed to authoctonous microorganisms or other contaminants, which are independent of human faecal pollution. One plausible mechanism is the exposure to other non-human pathogens, which would not be expected to correlate with human FIO concentrations. Avian schistosomes are parasitic trematodes shed in the faeces of waterfowl and can cause cercarial dermatitis (swimmer's itch) through direct skin penetration (Horák et al., 2015). Whilst clinical cases of cercarial dermatitis are sparsely reported in Hungary (Juhász et al., 2022), the true prevalence may be underestimated, as mild cases are likely to resolve spontaneously without prompting medical consultation. Alternatively, whilst our study sites were not directly affected by wastewater treatment effluents, agricultural or industrial runoffs could have contaminated the freshwater with chemical irritants.

An RCT design addresses the key methodological limitations of observational cohort studies in freshwaters by enabling robust, exposure-response based relative risk estimation, under controlled conditions. The Epibathe study thus addressed a critical evidence gap surrounding the use of FIOs to monitor freshwater quality. Wiedenmann et al. (2005) remains the only previously published RCT to have examined the health risks following controlled freshwater exposure with concurrent measurement of FIOs. Similar to our findings (Wiedenmann et al.), reported positive associations between increasing *E. coli* concentrations and GI illness risk. However, their No Observed Adverse Effect Level (NOAEL) framework relies upon a single exposure threshold, above which presents increased risk of illness: whilst readily applicable for regulatory purposes, this fails to quantify risk across the full exposure-response gradient. Instead, our analyses provide population-average relative risks across a broad range of FIO exposures, enabling a more detailed epidemiological characterisation of such exposure-response relationships.

In sites dominated by point-source sewage contamination, the association between FIO concentrations and the risk of GI illness is well established, forming the basis for existing regulatory standards (Cabelli et al., 1983; Office of Water Quality, 2012). In contrast, evidence for the health impact of non-point contamination, which is largely derived from cohort studies, has been inconsistent, resulting in uncertainty surrounding the predictive value of FIOs in such settings (Fujioka et al., 2015). With the superior study design of a RCT, our findings demonstrate statistically significant associations between GI illness and both *E. coli* and somatic coliphages across freshwater environments characterised by diffuse contamination. By combining site-clustered analyses across a variety of freshwater environments and a broad exposure gradient, this Epibathe trial robustly extends the conclusions of Wiedenmann et al. (2005): the consistency of findings across the different freshwater systems of Hungary and Germany respectively provides a strong empirical foundation for the wider applicability of these exposure-response relationships. Further studies will be necessary to confirm these findings in other climatic and environmental settings and thereby strengthen the epidemiological evidence base for specific FIOs, relevant to recreational water quality guidelines internationally.

Several limitations should be considered. The sample size was set with reference to GI illness, benchmarked against prior recreational bathing studies (Kay et al., 1994; Wiedenmann et al., 2005). Whilst the Epibathe study proved sufficiently powered to detect the GI illness-specific exposure-response associations for both *E. coli* and somatic coliphages, two key factors reduced the overall power actually achieved. Firstly, the prevalence of GI illness within the population were unexpectedly lower than those reported in earlier trials at comparable indicator densities. Secondly, no sample-size considerations were given for the non-enteric health outcomes, which proved to be considerably rarer than GI illness: respiratory (0.38%), ear (0.55%) and eye (0.55%) infections together contributed between nine and thirteen events each,

precluding robust inference. Accruing comparable numbers of events for these outcomes would require a cohort approximately five times larger, in the order of 13,000 participants, together with a greater number of study sites to capture a wider exposure gradient and improve cluster-level inference. Concurrent characterisation of individual-level exposure is logistically demanding and costly: these constraints account for the scarcity of controlled freshwater bathing trials, of which only Wiedenmann et al. (2005) preceded the present work.

Non-compliance with trial instructions could also introduce biases through exposure misclassification. Additional freshwater bathing outside the trial window was reported by 25 and 27% of non-bathers and bathers respectively (Table S7 in Supplementary Text 2). For the bather-versus-non-bather comparison, this would bias the exposure-response relationship towards the null, which is consistent with the non-significant association between bathing status and GI illness (Table 2). Crucially, this bias would not affect our primary exposure-response estimates, which are derived from within the bather group from zone- and time-specific FIO titres.

Finally, while FIO measurements were temporally proximate to faecal exposure, they remain indirect proxies for the presence of human enteric pathogens: our microbiological methodologies could not exclude the possibility that elevated FIO concentrations originated from non-human faecal contamination. These constraints highlight the potential value of RCTs with improved study designs, greater statistical power and environmental diversity, as well as the evaluation of a broader panel of FIOs, including source-specific markers, that may better reflect health risks to humans.

5. Conclusion

The Epibathe trial quantified GI illness risk across a broad range of FIO exposures in freshwater recreational environments. Our findings reinforce the value of *E. coli* as a FIO for GI illness and highlight the growing importance of somatic coliphages as indicators of bathing water quality, whilst providing evidence that intestinal enterococci have limited predictive value in freshwater contexts. These results are consistent with those from the only previously published similar RCT epidemiological study in freshwater (Wiedenmann et al., 2005).

CRediT authorship contribution statement

Edward K.S. Lam: Formal analysis, Software, Writing – original draft, Writing – review & editing. **Márta Vargha:** Data curation, Investigation, Methodology, Project administration, Resources, Writing – review & editing. **Daniel Thomas:** Investigation, Project administration. **Roland Salmon:** Investigation, Project administration. **Mihály Kádár:** Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Supervision. **Maria José Figueras Salvat:** Investigation, Methodology. **Julii Brainard:** Conceptualization, Project administration, Writing – review & editing. **Jamie Bartram:** Investigation, Methodology, Project administration. **Zsofia Barna:** Data curation, Investigation, Methodology, Project administration, Resources. **Paul R. Hunter:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing.

Data and code availability

Available at https://github.com/edwardkslam/Epibathe

Funding

This research was funded by the 6th Research Framework Program of the European Union (grant no. 022618), and the UK National Institute for Health Research (NIHR) Health Protection Research Unit in Gastrointestinal Infections, grant no. NIHR207399, in partnership with the UK Health Security Agency (UKHSA). Any views expressed in this article are those of the authors and not necessarily those of the funders, our employers, UKHSA or NIHR.

Declaration of competing interests

All authors declare no conflicts of interest.

Acknowledgements

We would like to acknowledge Prof David Kay, Professor Emeritus at Aberystwyth University, for his leadership in leading the Epibathe project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2026.114888.

References

Anderson, K.L., Whitlock, J.E., Harwood, V.J., 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl. Environ. Microbiol. 71 (6), 3041–3048.

Binns, K., 2023. When water pollution is rife in the UK, where's safe for wild swimming? from. https://www.nationalgeographic.com/travel/article/water-pollution-uk-where-wild-swimming.

Cabelli, V., Dufour, A., McCabe, L., Levin, M., 1982. Swimming-associated gastroenteritis and water quality. Am. J. Epidemiol. 115 (4), 606–616.

Cabelli, V., Dufour, A., McCabe, L., Levin, M., 1983. A marine recreational water quality criterion consistent with indicator concepts and risk analysis. Water Pollution Control Federation 55 (10), 1306–1314.

Cromley, V., Mackintosh, C., 2021. The economic impact of outdoor recreation in the UK: the evidence. from. https://researchonline.ljmu.ac.uk/id/eprint/1579/1/Reconomics%20-%20The%20Evidence%20PDF%202014.pdf.

Davies, C.M., 1989. The Effect of Sunlight and Other Factors on the Survival of Starved Enteric Bacteria in Natural Waters. Newcastle University.

Dufour, A.P., 1984. Health Effects Criteria for Fresh Recreational Waters. Environmental Protection Agency. August. from.

European Union, 2006. Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 Concerning the Management of Bathing Water Quality and Repealing Directive 76/160/EEC. Document 32006L0007. European Union, European Union. Document 32006L0007.

Ferley, J., Zmirou, D., Balducci, F., Baleux, B., Fera, P., Larbaigt, G., Jacq, E., Moissonnier, B., Blineau, A., Boudot, J., 1989. Epidemiological significance of microbiological pollution criteria for river recreational waters. Int. J. Epidemiol. 18 (1), 198–205.

Fewtrell, L., Kay, D., 2015. Recreational water and infection: a review of recent findings. Curr. Environ. Health Rep. 2 (1), 85–94.

Fózer, M., Nagy, M.T., Csépes, E., Cserháti, M., Kaszab, E., 2024. Potenciális szennyezők monitorozása a Közép-Tisza szakaszon vízhigiénés szempontból. Journal of Central European Green Innovation 12 (1–2), 66–75.

Fujioka, R., Narikawa, O., 1982. Effect of sunlight on enumeration of indicator bacteria under field conditions. Appl. Environ. Microbiol. 44 (2), 395–401.

Fujioka, R.S., Solo-Gabriele, H.M., Byappanahalli, M.N., Kirs, M., 2015. US recreational water quality criteria: a vision for the future. Int. J. Environ. Res. Publ. Health 12 (7).

Gascon, M., Zijlema, W., Vert, C., White, M.P., Nieuwenhuijsen, M.J., 2017. Outdoor blue spaces, human health and well-being: a systematic review of quantitative studies. Int. J. Hyg Environ. Health 220 (8), 1207–1221.

Geneshka, M., Coventry, P., Cruz, J., Gilbody, S., 2021. Relationship between green and blue spaces with mental and physical health: a systematic review of longitudinal observational studies. Int. J. Environ. Res. Publ. Health 18 (17), 9010.

Hall, V., Taye, A., Walsh, B., Maguire, H., Dave, J., Wright, A., Anderson, C., Crook, P., 2017. A large outbreak of gastrointestinal illness at an open-water swimming event in the River Thames, London. Epidemiol. Infect. 145 (6), 1246–1255.

Horák, P., Mikeš, L., Lichtenbergová, L., Skála, V., Soldánová, M., Brant, S.V., 2015. Avian schistosomes and outbreaks of cercarial dermatitis. Clin. Microbiol. Rev. 28 (1), 165–190.

Huber, P.J., 1967. The behavior of maximum likelihood estimates under nonstandard conditions. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, CA.

Juhász, A., Majoros, G., Cech, G., 2022. Threat of cercarial dermatitis in Hungary: a first report of Trichobilharzia franki from the mallard (Anas platyrhynchos) and European ear snail (Radix auricularia) using molecular methods. Int. J. Parasitol.: Parasites Wildl. 18, 92–100.

Kay, D., Jones, F., Wyer, M., Fleisher, J., Salmon, R., Godfree, A., Zelenauch-Jacquotte, A., Shore, R., 1994. Predicting likelihood of gastroenteritis from sea bathing: results from randomised exposure. Lancet 344 (8927), 905–909.

Korajkic, A., McMinn, B.R., Harwood, V.J., Shanks, O.C., Fout, G.S., Ashbolt, N.J., 2013. Differential decay of enterococci and Escherichia coli originating from two fecal pollution sources. Appl. Environ. Microbiol. 79 (7), 2488–2492.

Kozak, S., Roiko, A., Gutjahr-Holland, K., Ahmed, W., Veal, C., Fisher, P., Toze, S., Weir, M., Stratton, H., 2025. The use of faecal indicator organisms to manage microbial health risks in recreational waterways not impacted by point sources of sewage: a systematic review of the epidemiological evidence. J. Water Health 23 (5), 563–586.

Liang, K.-Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. Biometrika 73 (1), 13–22.

Office of Water Quality, 2012. Recreational water quality criteria. from. https://www.epa.gov/sites/default/files/2015-10/documents/rwqc2012.pdf.

Rubin, D.B., 1996. Multiple imputation after 18+ years. J. Am. Stat. Assoc. 91 (434), 473–489.

Rusiñol, M., Fernandez-Cassi, X., Hundesa, A., Vieira, C., Kern, A., Eriksson, I., Ziros, P., Kay, D., Miagostovich, M., Vargha, M., 2014. Application of human and animal viral microbial source tracking tools in fresh and marine waters from five different geographical areas. Water Res. 59, 119–129.

RYA, 2023. Watersports participation remains buoyant according to latest Watersports Participation Survey. from. https://www.rya.org.uk/news/watersports-participation-remains-buoyant-2023.

University of Brighton, Exegesis SDM Ltd, Rubicon Associates, Plumpton College and G. a. L. Hughes, 2015. Enjoying water. from. https://assets.publishing.service.gov.uk/media/5a7ba5f440f0b638d61be082/gean0609bqda-e-e.pdf.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica: J. Econom. Soc. 817–838.

White, I.R., Royston, P., Wood, A.M., 2011. Multiple imputation using chained equations: issues and guidance for practice. Stat. Med. 30 (4), 377–399.

White, M.P., Elliott, L.R., Gascon, M., Roberts, B., Fleming, L.E., 2020. Blue space, health and well-being: a narrative overview and synthesis of potential benefits. Environ. Res. 191, 110169.

WHO, 2021. Guidelines on recreational water quality. In: Coastal and Fresh Waters, ume 1. World Health Organization.

Wiedenmann, A., Krüger, P., Dietz, K., López-Pila, J.M., Szewzyk, R., Botzenhart, K., 2005. A randomized controlled trial assessing infectious disease risks from bathing in fresh recreational waters in relation to the concentration of Escherichia coli, intestinal enterococci, Clostridium perfringens, and somatic coliphages. Environ. Health Perspect. 114 (2), 228.

World Health Organization, 2011. Guidelines for drinking-water quality. from. https://iris.who.int/server/api/core/bitstreams/b437749a-43f7-472f-a6d9-42596e8ac0ae/content.

World Health Organization, 2021. Guidelines on recreational water quality. In: Coastal and Fresh Waters, vol. 1. World Health Organization.