ACS ES&T Water
pubs.acs.org/estwater

Article

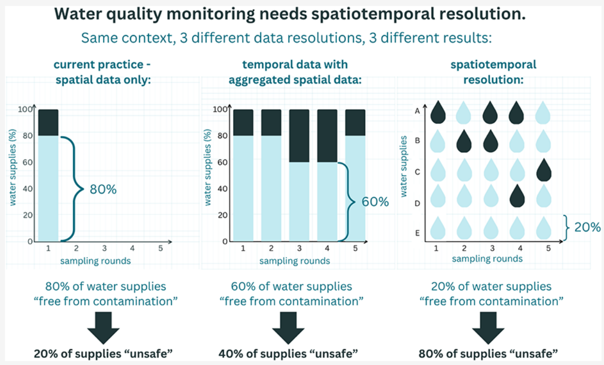
# Intermittent Contamination of Small Water Supplies: Implications for Effective Monitoring and JMP Global Assessment

Raphaela Betz, Saskia Nowicki, Edith Darin, Jamie Bartram, and Katrina J. Charles*

Cite This: https://doi.org/10.1021/acsestwater.6c00247 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Access to "safe" drinking water is often inferred from single-point measurements or isolated cross-sectional surveys, which do not reflect temporal variability. We examine spatiotemporal variability in *Escherichia coli* contamination in water systems and evaluate implications for water quality monitoring. Using sequence analysis, a method to evaluate ordered sets of data, we assessed patterns of contamination from longitudinal monitoring of 524 small water supplies (8–12 rounds over 1 year) in Bangladesh, Nepal, and Tanzania. The analysis demonstrated that detectable fecal contamination of drinking water is an intermittent condition rather than a stable condition in up to 95% of small water systems. Isolated cross-sectional surveys underestimated the period prevalence of contaminated sources by factors of 1.4–5. Sampling in wetter weather conditions improved detection of at-risk sources. These findings suggest substantive overestimation of safety in system-specific and national monitoring and assessments of global status and progress toward the Sustainable Development Goals target. For small water systems, we propose a shift to a risk-based operational monitoring approach that prioritizes timing over frequency for efficiency of hazard identification and resource use.




**KEYWORDS:** *drinking water safety, E. coli, microbial contamination, surveillance, water quality monitoring*

## 1. INTRODUCTION

Access to safe drinking water is fundamental for protecting public health and a core aim of Sustainable Development Goal (SDG) 6.1. Yet an estimated 2.1 billion people globally still lack access to safely managed drinking water services, and recent modeling suggests that this number may exceed 4.4 billion in low- and middle-income countries.[1,2] In 2021 alone, unsafe water was linked to over 800,000 deaths and 41 million disability-adjusted life years (DALYs), with the burden falling disproportionately on children under five in sub-Saharan Africa and South Asia.[3] A major driver of this disease burden is exposure to fecal contamination, a major cause of acute diarrhea and a leading cause of death in children under five, as well as long-term effects such as stunting.[4] It has been suggested that even occasional exposure to contaminated water can be harmful: Hunter et al.[5] calculated that just a few days of drinking unsafe water can effectively erase the health benefits of safe drinking water access.

Controlling the health risks from fecal contamination requires that drinking water be not only contamination-free at the moment of sampling—but also consistently safe over time. The WHO Guidelines for Drinking-Water Quality reflect this principle by stating that water must be "consistently free from fecal contamination" to be considered safe.[6] Despite this, most national and international water quality monitoring programs—including the Multiple Indicator Cluster Surveys (MICS) and many national regulatory frameworks—rely on isolated cross-sectional sampling.[7] These programs assess drinking water safety by taking microbial samples from a selection of locations at one point in time and are often infrequent, typically undertaken every few years. For small water systems, there may be little to no other water quality monitoring. While cross-sectional snapshots offer insights into spatial variation, they lack the temporal dimension needed to assess whether individual sources provide consistently uncontaminated, and therefore safe, water. As Charles et al.[8] note, there is a critical difference between a *sample*, which reflects conditions at a single moment, and *safety*, which requires stable and reliable performance over time. Cross-sectional sampling is appropriate for chronic conditions but will structurally underestimate conditions that are intermittent; e.g., in studies of malaria prevalence, parasite densities will fluctuate to below detectable levels,[9] with single samples

Received: February 19, 2026
Revised: July 15, 2026
Accepted: July 16, 2026

known to misclassify risk.[10] For water safety, we can consider this in terms of the prevalence-incidence (Neyman) bias:[11] counting only contaminated water systems at one point in time (prevalence) underestimates water safety over a period of time (incidence), as contamination is not only driven by persistent sources but also by short-duration spikes.

Evidence from longitudinal studies demonstrates the variability in bacteriological contamination of drinking water. Seasonal trends are the best-studied aspect of temporal variability, with most studies reporting greater contamination during rainy periods.[12−18] Rainfall and temperature influence water quality across a range of time scales, from short-term hourly shifts to interannual climate-driven changes.[19−25] These shifts in water quality are also reflected in health outcomes, with seasonal variations in the impact of water interventions on diarrhea.[26] Beyond seasonality, studies have noted temporal variability related to human factors (e.g., infrastructure or population density) and environmental factors (e.g., geology)[27−29] as well as imprecision and inaccuracy in available test methods.[30,31] This confirmed variability indicates that relying on cross-sectional sampling may lead to misclassification of intermittently contaminated sources as safe, thereby underestimating the health risk and overestimating progress toward national and SDG targets.

For intermittent contamination, a different monitoring approach is required. Recommended sampling frequencies for small water systems vary, without an evidentiary basis on the risks captured. WHO recommends monitoring for *E. coli* in small water systems monthly.[32] By contrast, in Europe, the Drinking Water Directive[33] specifies sampling frequencies of at least once per year for systems producing less than 10 $m^3$ per day and twice per year for systems up to 100 $m^3$ per day. Small water systems typically produce low volumes, for example, 1.4 $m^3$ per day for a hand pump and 1.8 $m^3$ per day for an electric pump kiosk in Mali.[34] Sampling monthly is costly, especially where systems are located in rural areas.[35] The challenge arising is that sampling is needed to move toward expanding provision of consistently safe drinking water, but sampling once a year may not capture intermittent contamination and may misdirect resource use.

Our aim in this paper is therefore to provide evidence on the spatiotemporal patterns in drinking water contamination to inform better design of risk-based sampling approaches to understand water safety and to enhance monitoring to inform water safety improvements. Using secondary data from longitudinal sampling programs in Bangladesh, Nepal, and Tanzania, we analyze spatiotemporal contamination dynamics down to the individual water point level based on 5,020 *E. coli* samples from 524 points of collection (PoCs). We apply sequence analysis, a method for evaluating ordered sets of longitudinal states, to analyze and visualize variability across space and time. This allows us to examine how often intermittently contaminated sources are missed by isolated cross-sectional sampling, which is used to explore the implications for our understanding of drinking water safety globally. Based on our findings and the limitations of cross-sectional sampling for understanding intermittent contamination, we propose a focus on operational monitoring, providing practical steps for designing risk-based microbial water quality monitoring to support progress toward consistently safe water.

## 2. MATERIALS AND METHODS

### 2.1. Longitudinal Drinking-Water Data

We utilized secondary data from longitudinal sampling of water systems designed to understand the impact of weather on access to water, sanitation, and hygiene.[19] Briefly, study locations across three countries were selected with the intention of capturing how the performance of drinking water supply services varies with weather, with sites selected for recognized vulnerability to climate hazards. Within selected sites, sampling was focused on improved water supplies, with other sources included if nearby households were using them. Outcomes for power calculations were focused on household drinking water, which is not evaluated here. Water quality was sampled from the PoC up to 16 times over at least 12 months. The duration and the intervals of the sampling rounds varied. Samples were taken at the PoC, after flaming or sterilization, so results represent the water "as delivered",[36] excluding contamination from hygiene conditions at the tap. 100 mL water samples were transported to appropriate laboratories or dedicated field laboratories, maintaining the cold chain, and analyzed for *E. coli* using membrane filtration methods (see Supporting Information 1). For the analysis presented here, the inclusion criteria selected for consistent and comparable data sets: (1) sampling rounds were selected if timings were comparable across multiple locations within a country to reflect distribution across seasons in different climate contexts (see Supporting Information 1 for timing and weather conditions), with 8/9 rounds included for Tanzania, 8/8 for Nepal, and 12/16 for Bangladesh; and (2) individual PoCs had *E. coli* test results available for all included sampling rounds to reflect variability. Seasonal water sources were not included. 137 individual water supplies (1,096 samples) for Tanzania, 178 supplies (1,424 samples) for Nepal, and 207 supplies (2,484 samples) for Bangladesh were included in the analysis. In Tanzania, two samples were taken per PoC per sampling round, with the mean used in this analysis.

### 2.2. Analysis of Patterns of *E. coli* Detection

We differentiated between a contaminated sample, i.e., where *E. coli* was detected, and a contaminated PoC, where *E. coli* was present in at least one sample over the duration of the study. Sequence analysis was used to characterize spatiotemporal *E. coli* contamination, including intraround variability and longitudinal patterns of individual PoCs. Sequence analysis, increasingly used in the social sciences,[37] examines ordered sets of states over time. In this analysis, each sequence represents the full set of sampling results from one PoC. *E. coli* concentrations were converted into the four WHO risk categories—low (<1 cfu/100 mL), intermediate (1−10), high (11−100), and very high (>100)—to provide ordinal input for analysis. While this transformation reduces numerical granularity, it mitigates threshold inconsistencies and enhances cross-country comparability.[38,39]

Analyses were performed in R using TraMineR,[40,41] and are available in an open repository (see Supporting Information 2). Spatiotemporal distributions of risk levels were visualized with modal state plots and sorted sequence index plots (SIPs). SIPs were generated through *optimal matching*, which orders sequences by their divergence from an ideal state (here: continuous "low risk"). They provide a compact depiction of within-source and between-round variability. The seqstatd function was used to summarize the frequency of unique sequence patterns by country.

Complementary analyses employed the tidyverse suite. For each PoC, the *E. coli* detection frequency (the proportion of sampling rounds with detectable contamination) was calculated, and distributions were visualized using stacked column plots. Relationships between detection frequency and maximum *E. coli* concentration were visualized with scatter plots. The most effective (most contaminated PoCs detected) and least effective (fewest contaminated PoCs detected) sampling combinations were identified by calculating, for each incremental increase in the number of samples, which combination of sampling rounds had the highest detection rate of contaminated PoCs and which had the lowest. The "effectiveness" of a set of sampling rounds to reflect the water quality of a single PoC was

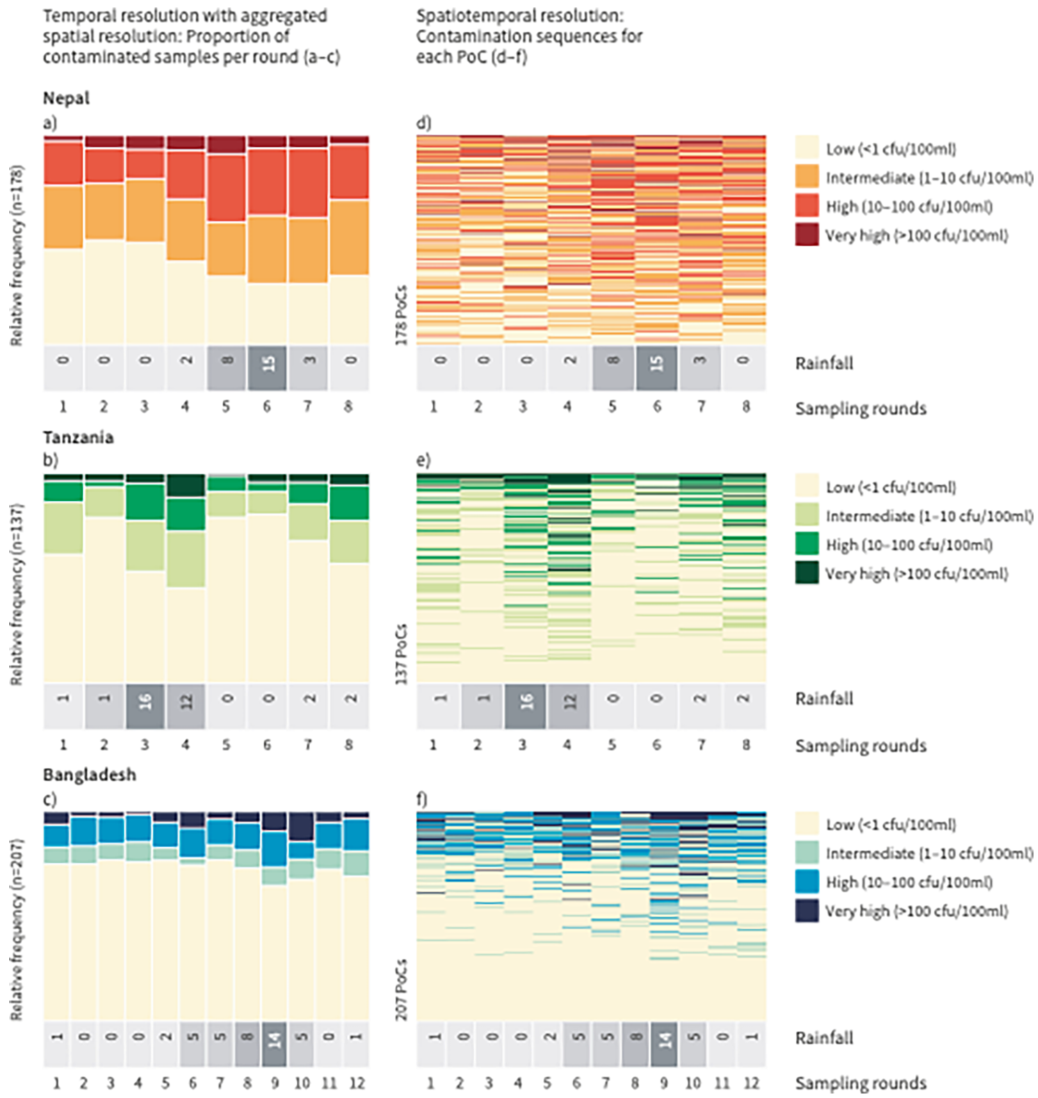



**Figure 1.** Variability in *E. coli* contamination in drinking water: (a−c) the frequency of contaminated samples per sampling round; (d−f) the variability in patterns of contamination for individual PoCs presented as sorted sequence index plots (each horizontal line in the graphs shows the contamination sequence of one individual PoC, with PoCs sorted from most contaminated (top) to least contaminated (bottom)). Country data are arranged from the context with the most overall contamination (Nepal) to the context with the least overall contamination (Bangladesh). Gray bars provide daily rainfall (mm), averaged across the sampling period across sites within each country.

represented through the incremental likelihood of detection across different combinations and numbers of sampling rounds, which was calculated using the "combn" function. Results were compared with rainfall patterns reported by Charles et al.[19] (see Supporting Information 3), using relative rankings of average daily rainfall during the sampling round within each country. We have avoided categorizing by wet/dry season to reflect the extremes that occur within these seasons. This also helps to provide operational definitions of rainfall that can inform sampling design in diverse contexts.

Data on the three countries were analyzed separately because of differences in sampling rounds and climate. We were interested in the patterns of intermittently detectable contamination, which was expected to be driven by weather and highly localized factors (e.g., geology, quality of construction and maintenance, local pollution sources) rather than the general categorization of PoC type (e.g., piped water, borehole with hand pump, etc.). Hence, within countries, the PoC types were aggregated.

## 2.3. Comparison to Global Water Quality Data Sets

Data from nationally (and regionally) representative studies of drinking water quality at the PoC, employing single-sample cross-sectional methods, provided a comparison to explore the implications of the effectiveness of individual sampling rounds. PoC *E. coli* data were used from the sixth round of UNICEF Multiple Indicator Cluster Surveys (MICS6), where water quality data had been collected from 34 surveys across 31 countries (Pakistan included 4 provinces separately). Data sets including water quality data were retrieved from the MICS website. The MICS6 survey water quality sampling method at the PoC requires flushing of the PoC but not cleaning, providing an estimate of water quality "as collected", inclusive of localized contamination on the tap or spout.[36] Data were used unweighted, representing samples rather than population, as this is comparable with the analysis of longitudinal data above. To extrapolate the potential underestimation by the WHO/UNICEF Joint Monitoring Programme (JMP) due to interpretation of risk from single samples, three comparison points were selected from analysis of the effectiveness of sampling rounds: the most effective single sampling round in any one country (i.e., the highest proportion of contaminated PoCs detected), the least effective single sampling round, and an intermediate round. The proportions of PoCs sampled that were contaminated in the MICS data (see Supporting Information 4) and the proportion that would be expected to be contaminated based on the effectiveness of sampling were calculated.



https://doi.org/10.1021/acsestwater.6c00247
ACS EST Water XXXX, XXX, XXX−XXX

## 3. RESULTS AND DISCUSSION

### 3.1. Spatiotemporal Variability in Contamination

Spatiotemporal variability of *E. coli* contamination in drinking water samples demonstrates intermittent detectable contamination in the majority of settings, with strong seasonality (Figure 1). For the data set used, seasonal variability in *E. coli* contamination was previously reported,[19] demonstrating significant positive correlations between rainfall and contamination of improved water sources in most cases, with exceptions for piped water in Bangladesh and protected wells in two sites in Tanzania. The subset of data used in this study confirms seasonal variability in the proportion of contaminated samples in spatially aggregated data across sampling rounds (Figure 1a−c). The magnitude of this variability differs between countries: in Bangladesh, *E. coli* detection was highest in sampling round 9 (July), coinciding with the heaviest rainfall, with the most "very high" risk results toward the end of the monsoon in September (round 10); in Tanzania, rates of contamination were higher than in Bangladesh, with two peaks coinciding with rainfall, with higher rainfall in April−May (rounds 3−4) and lower rainfall in rounds 7−8; in Nepal, high rates of contamination were observed throughout the sampling period, with an increase over the rainy months from July to November (rounds 6−8) (see Supporting Informations 1 and 3 for more details on weather conditions).

The sequences of contamination for individual PoCs (Figure 1d-f) are more complex than these broader seasonal patterns. In these Sequence Index Plots (SIPs), sorting is based on frequency and intensity of risk. They show that contamination at the individual PoC, each presented as an individual horizontal row, is intermittent rather than following a particular pattern. We adopt the term "intermittent" to represent the irregularity demonstrated by these SIPs. The SIPs are sorted from PoCs with the most frequent and highest risk levels (top, darker colors) to the lowest risk levels (bottom, lightest colors).

These SIPs highlight three important aspects of the contamination.

*First, small water supplies are not providing consistently safe drinking water.* Fecal contamination was detected at least once during the sampling campaigns in the majority of PoCs: 69% in Bangladesh, 91% in Tanzania, and 99% in Nepal (Figure 2). In Nepal, 77% of PoCs were contaminated at least half of the time during the sampling occasions. This was true for 32% of PoCs in Tanzania and 24% of PoCs in Bangladesh (see Supporting Information 3). The majority of the water systems sampled were boreholes, tubewells with hand pumps, or piped systems with no regular treatment before the PoC.[19] This lack of treatment is globally widespread, as reflected by PoC water quality in cross-sectional sampling.[7] From the spatially aggregated data, which reflects what we would observe from cross-sectional sampling (Figure 1a-c), the proportion of PoCs with *E. coli* detected in a single sampling round ranged from less than 20% (Tanzania) to around 70% (Nepal) across the three countries.

*Second, contamination in small water supplies is intermittent.* The SIPs demonstrate that contamination is neither consistent nor does it follow the seasonal patterns observed in the aggregated data. This pattern also applies to the level of contamination: where contamination is generally high, there are occasional samples that indicate low risk, and vice versa. Water supplies that are generally free from *E. coli* occasionally

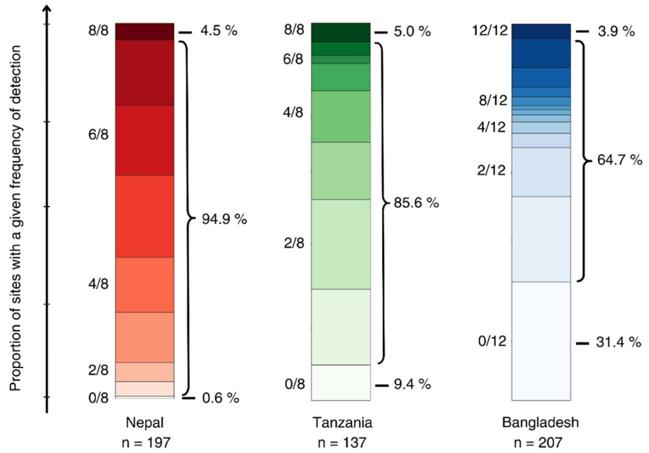



**Figure 2.** Frequency of *E. coli* detection at individual PoCs. This chart shows the percentage of water sources that experienced different frequencies of contamination across sampling rounds. "0/8" or "0/12" indicates PoCs that were consistently free of *E. coli*, while "8/8" or "12/12" denotes PoCs that were contaminated in every round. The color gradient represents the frequency of *E. coli* detection (presence/absence).

have a sample indicating high or even very high risk. For all three countries, only 4−5% of PoCs were consistently contaminated, and most PoCs were inconsistently contaminated (65% for Bangladesh, 86% for Tanzania, 95% for Nepal). Some of this variability may be attributable to the challenge of monitoring *E. coli*, particularly in more rural areas, associated with cross-contamination during sampling, maintaining the cold chain during travel, and clumping of *E. coli* in analysis. However, variability is well supported by the wider literature and has considerable implications for the design of risk-based and cost-effective monitoring programs.

*Lastly, there were no shared patterns of contamination sequences.* Contamination sequences varied considerably across PoCs, with no trends emerging. Contamination was more common in wetter periods, but the multiple sampling occasions during rainy seasons identified variability within those rainy periods, such as between the onset of rains and sustained rains (see Supporting Information 1 for more details on weather conditions). Unique patterns were common, observed in 49% of PoCs in Bangladesh, 60% in Tanzania, and 94% in Nepal (see Supporting Information 3). The higher proportion of shared contamination patterns in Bangladesh and Tanzania is largely due to consistently noncontaminated PoCs; Nepal, having the highest rate of contamination, had the strongest sequence variety, with only five patterns appearing twice, and 94% (168 PoCs) exhibiting unique patterns. These findings highlight the complexity of contamination dynamics, where each PoC may follow a specific pattern, complicating efforts to generalize or predict contamination solely based on trends in aggregated data. These differences may be driven by different contamination pathways. For example, rapid contamination pathways such as surface water ingress into a poorly sealed borehole occurring immediately after rain would have a different pattern compared to a borehole where contamination is via a leaking sewer with contamination moving through deeper groundwater. Not all contamination pathways will be equally affected by rainfall.

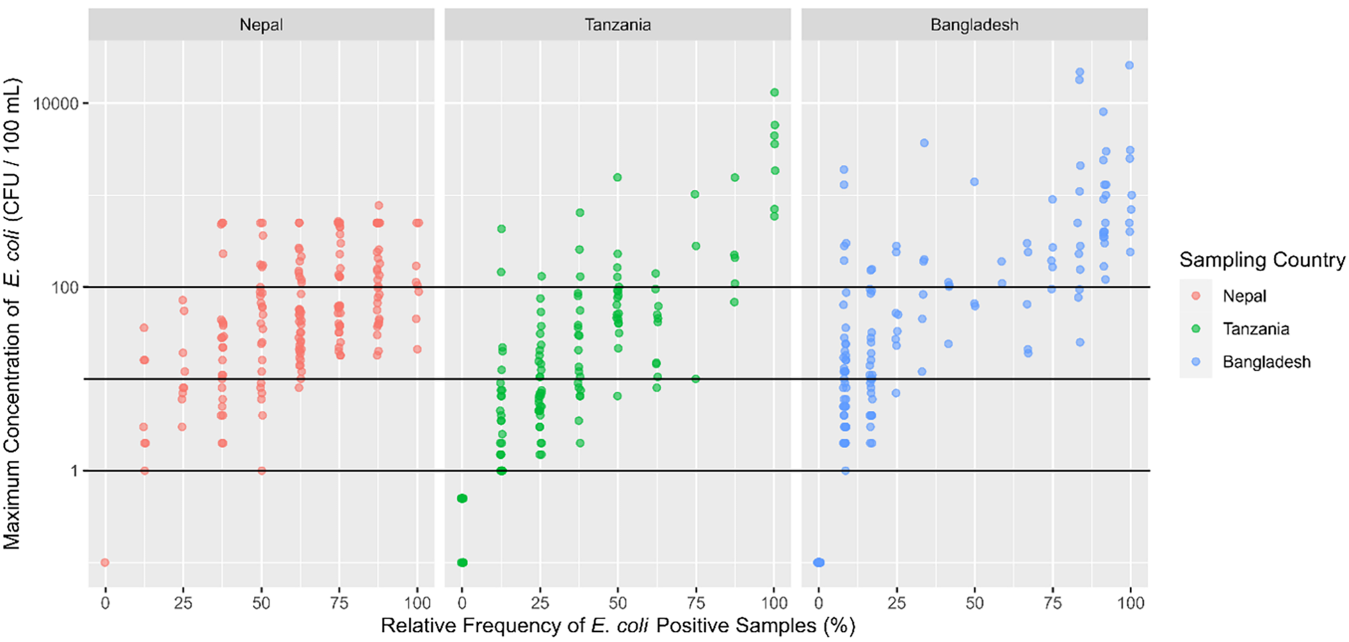



**Figure 3.** Maximum *E. coli* concentration observed for each water supply compared to the frequency of *E. coli*-positive samples for each individual PoC. The maximum detection limit in Nepal was 200 cfu/100 mL, and Tanzania data are based on the mean of two tests. The black horizontal lines indicate the cutoffs between risk categories used by WHO.

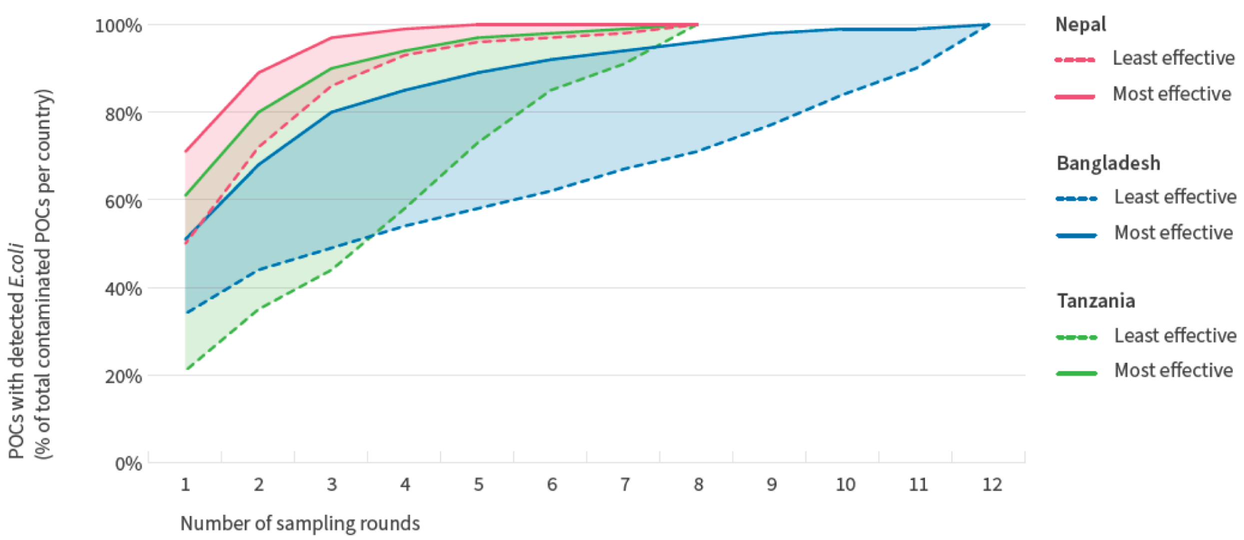



**Figure 4.** Model of contaminated PoCs where *E. coli* would have been detected based on cumulative effectiveness of sampling rounds. The solid lines show the sampling-round combination with the highest percentage of identified contaminated PoCs. The dotted lines show the sampling round combination with the lowest percentage of detection. Accordingly, the detection results of all other combinations fall within the shaded area.

### 3.2. Relationship between Severity and Frequency

More severe contamination (higher *E. coli* concentration) was associated with a higher frequency of contamination (Figure 3). This was a common trend in the data but not universal. Samples in the intermediate risk range (1−10 cfu/100 mL) were more common in systems where contamination was detected 25 to 50% of the time. For all sources where contamination was detected less than 50% of the time, the majority of sampling results are below 100 *E. coli*/ 100 mL, i.e., at risk levels "intermediate" or "high". For those PoCs which were contaminated 50% of the time or more, levels of contamination were highest. Even "very high" risk-level contamination was found in PoCs that were only contaminated once, albeit rarely. This supports the interpretation of the WHO risk levels,[32] that detections at higher levels are more significant in terms of health risk. However, the limitations of *E. coli* as an indicator should be noted, specifically that pathogens may be present in the absence of *E. coli* and diverse pathogens may be present in the presence of *E. coli*.[8] This effect is likely reinforced by analysis limitations, with higher concentrations more readily measured.

### 3.3. Cumulative Measures: Understanding Water Safety through Multiple Rounds

We explored how effective sampling rounds are in understanding the prevalence of contaminated systems, considering the cumulative effect and differences in timing. We draw from epidemiology to use the term prevalence, recognizing that this is period prevalence, as the definition of a system as




https://doi.org/10.1021/acsestwater.6c00247
ACS EST Water XXXX, XXX, XXX−XXX

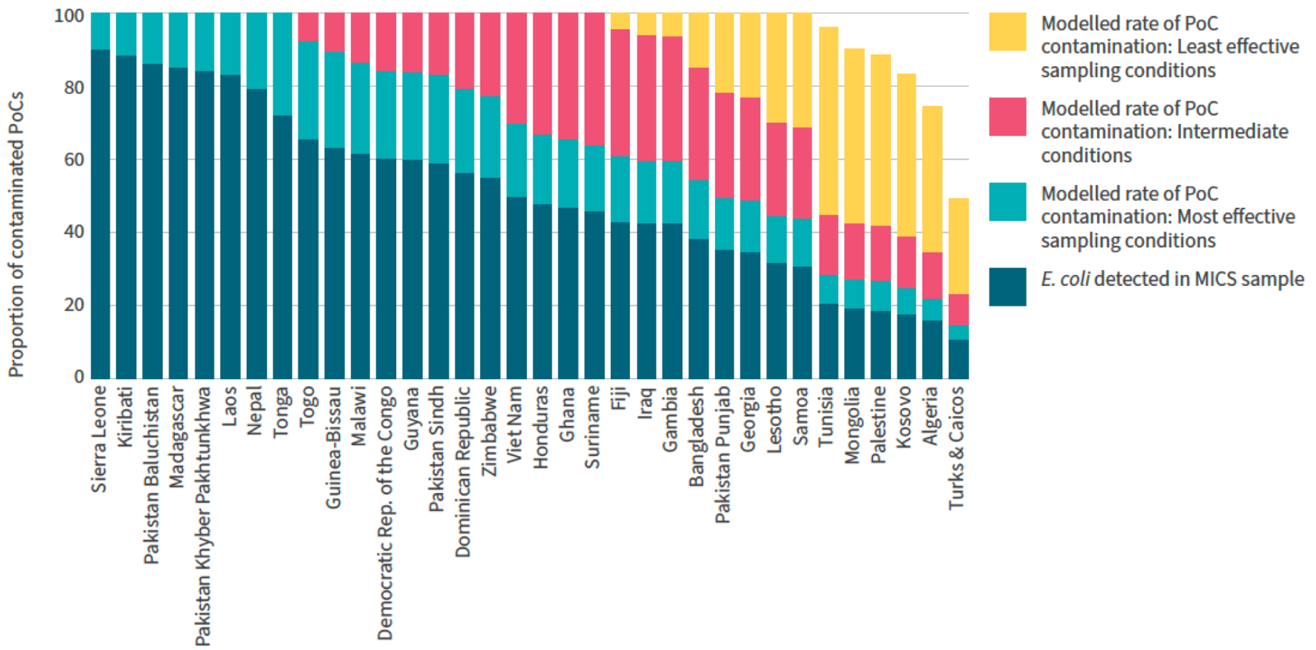



**Figure 5.** Estimated rates of contaminated PoCs: comparison of isolated cross-sectional water quality sampling campaigns by MICS, which reflect the proportion of PoCs with detected contamination, with modeled estimates of the period prevalence of unsafe water systems based on the effectiveness of sampling conditions.

contaminated is based on assessments over 12 months. The cumulative proportion of PoCs where contamination was identified increased with additional sampling rounds (Figure 4). The "most effective" rounds are those where more sampling highlighted risks (i.e., detected *E. coli*) in more of the contaminated PoCs. Where sampling focuses on the most effective rounds or conditions, the cumulative understanding of risks is nonlinear, with fewer samples able to support identification of the majority of unsafe PoCs. The difference between the most and least effective rounds is greater in Tanzania, where 90% of contaminated PoCs were identified in the three most effective sampling rounds, but only 44% in the three least effective; for Bangladesh, it was 80% versus 49%.

Comparing the three countries highlights commonalities from which other contexts can learn. In Nepal, where overall contamination is the highest, the timing of sampling appears to be less impactful, with two samples in less effective sampling periods matching the detection level of one sample in the most effective sampling period, and the minimum and maximum detection levels converging more quickly than those in the other two countries. In Bangladesh, where there is a similar monsoon rainfall pattern, the detection level in the most effective sampling round is as high as the combination of 3−4 sampling rounds in the least effective periods. By contrast, in Tanzania, which has a different climate but also high rates of contaminated PoCs, timing becomes very important, with four samples in the least effective periods needed to reach approximately the same detection level as that of one sample in the most effective period. As water safety improves, the timing of water sampling will become more critical to adequately capture risk, but weather will continue to play a key role.

Effective sampling combinations were considered to identify how to better capture intermittent contamination. Rainy periods were more effective sampling periods, and the lowest risk was in the driest periods; however, combining samples in wetter and drier periods captures different risks. In Nepal, the optimal two-round combination coincided with the rainiest periods (Table S3.1 in Supporting Information 3). When adding a third round, the best combination includes a drier period that saw a few isolated rain days in one of the two sites. The worst combination is adjacent sampling rounds in the driest period. Similar patterns were observed in Tanzania and Bangladesh, with wetter periods combined with drier periods to improve the effectiveness of sampling after two rounds. This approach aligns with risk-based approaches to sampling to understand different contamination pathways.

### 3.4. Implications for Global Estimates of Water Safety

Isolated cross-sectional samples contribute substantially to estimates of global access to safe drinking water. Using our analysis of how effective different sampling periods are in reflecting the period prevalence of contaminated water systems, we calculated the potential underestimation of prevalence by national reporting from isolated cross-sectional sampling with MICS. The data sets are not directly comparable to the sites included above, with differences expected by system type, climate as well as many other factors. The comparison is intended to provide a first estimation indicating the underestimation of the prevalence of unsafe water systems that may be expected in established national estimates, not a definitive measure. We used three reference points: (a) most effective detection rate of 71%: based on the most effective single cross-sectional sampling round in any one country in our data set, which identified 71% of the total PoCs that were contaminated at any point during the study period (as observed in Nepal); (b) Least effective detection rate of 21%: based on the least effective single cross-sectional sampling round, which identified 21% of the PoCs that were contaminated at any point in the study period (as observed in Tanzania); (c) and Intermediate detection rate of 45% (see Tables S3.1−3 in Supporting Information 3). For example, if a country was reported to have *E. coli* detected in 20% of water

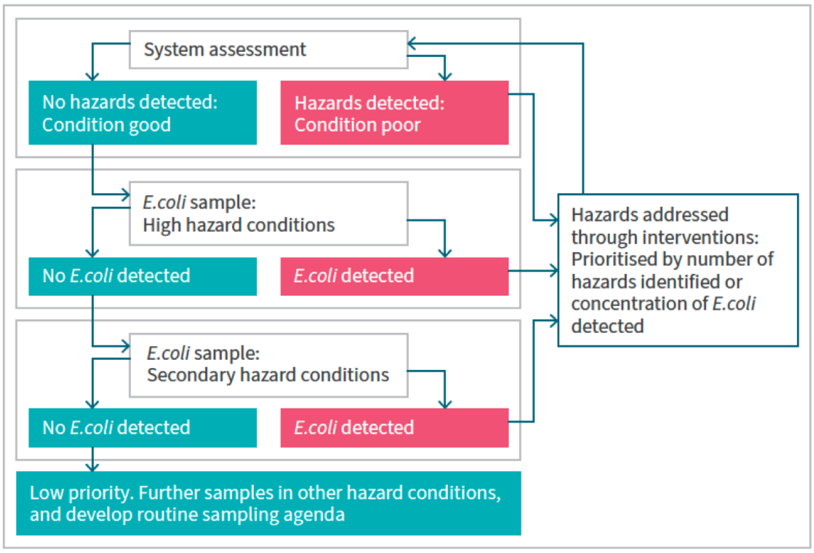



**Figure 6.** A risk-based monitoring design for operational monitoring in small water systems. Rainy seasons are often high-hazard conditions.

supplies at the PoC in an isolated cross-sectional survey, such as in MICS data from Tunisia, then, using our survey's most effective detection rate, this observed prevalence is expected to represent 71% of the total period prevalence. The period prevalence is therefore estimated at 28% (20% divided by 71%). Conversely, under the least effective sampling conditions, the period prevalence would be 96% (20% divided by the least effective detection rate 21%). In Figure 5, we present estimates for the expected proportion of contaminated PoCs based on these scenarios. This modeling indicates that in many countries or provinces it would be expected that all PoCs are not consistently free from fecal contamination. Even assuming that the most effective sampling conditions were present, 8 of the countries or provinces would be expected to have all PoCs contaminated. 20 of 34 countries and provinces (59%) are estimated to have no drinking water systems consistently free from fecal contamination based on our intermediate conditions. If the sampling had been done in the least effective sampling periods, as would be expected to be common due to the relative ease of sampling in dry compared to wet periods,[7] 28 of 34 countries and provinces (82%) are estimated to have no water supplies consistently free from fecal contamination.

These data resonate with the findings of Greenwood et al.,[2] that fecal contamination is more common than JMP estimates would indicate. There are considerable limitations to this modeling, and it is intended to indicate the potential scale of underreporting rather than to assign accurate measures of the prevalence of contaminated water systems. Notable limitations include that MICS data are nationally representative, including sites where water treatment is used, whereas the study sites where data were collected for this analysis were from areas facing different climate hazards and where water treatment was not common or reliable. In our data, as only PoCs with data available across all rounds were included, those that were seasonal were excluded. In each country, sites were selected that faced different hazards, including drought and flood, so the weather-related impacts may be overrepresented compared to national averages. We observe similarities between our results and the national results: Nepal has a high level of contamination in both our study (71% in the highest single round of sampling) and MICS (75%[42]), whereas Bangladesh has comparatively lower levels of contamination in both (35% in the highest single round in our study; 39% nationally in MICS[43]). Tanzania does not have MICS water quality data available. Lastly, detection of contamination would be expected to be lower in our data than in MICS due to different sampling protocols: MICS samples without cleaning the tap, whereas the protocol for these samples was to clean the tap to remove hygiene-related contamination.[36]

## 4. IMPLICATIONS FOR THE DESIGN OF WATER QUALITY MONITORING

The analysis has demonstrated that detectable fecal contamination of drinking water is an intermittent condition rather than a stable condition in untreated small water systems. It supports the widely understood principle of prevalence-incidence bias[11] that isolated cross-sectional sampling will structurally underestimate the prevalence of contaminated systems in these conditions. Similarly, the once-a-year approach used in Europe would underestimate the period prevalence of contamination in untreated systems but may be appropriate to verify that treatment and other water quality controls are effective. WHO recommendations for monthly monitoring would be expected to capture the hazards identified in this analysis; however, for small systems, the cost associated with these measures is often considered to be prohibitive.

Irregular patterns of contamination present a challenge for designing cost-effective water quality monitoring. Rather than advancing recommendations to improve cross-sectional sampling, we consider how our results can inform the development of cost-effective operational monitoring in small systems, particularly those without consistent and demonstrably effective treatment. This approach, summarized in Figure 6, aims to identify and understand hazards and prioritize actions in line with water safety planning approaches.[32]

We consider operational monitoring to be the first priority to improve water safety and can support incentive-based regulatory approaches.[44] Independent regulatory compliance monitoring ("surveillance") design may be informed by such




https://doi.org/10.1021/acsestwater.6c00247
ACS EST Water XXXX, XXX, XXX−XXX

approaches but will not advance water safety without operational management of water safety.

Testing for *E. coli* should not be the only or first water safety action. *E. coli* monitoring should be used to validate that existing water safety controls are performing effectively. The results presented highlight that contamination occurs intermittently and not in predictable patterns, reinforcing that a single sample in which *E. coli* is not detected does not identify a safe drinking water supply. Before investments in water quality testing, work should be undertaken to address the known hazards. Sanitary inspections,[32] with their appropriate corrective actions, or implementation and management of treatment should be considered before water quality monitoring.

Once *E. coli* have been detected, repeated testing does not provide new information unless actions have been taken to improve water safety, such as protecting the system or treating the water. Resources should be focused on reducing hazards and not on collecting data. After actions have been taken, another test can provide validation that the control measures have worked. We would equally recommend similar approaches for chemical contamination; e.g., once arsenic has been detected at a level indicating concern, the priority is to act to reduce arsenic exposure—further sampling is not the priority. If other measures identify a system as vulnerable to fecal contamination because controls are not operating sufficiently, such as insufficient chlorine residuals or high sanitary inspection hazards, the system is already at risk and priority should be placed on improvements; *E. coli* samples are not needed to demonstrate the risk.

Where water safety actions have been implemented and systems are appropriately managed to deliver safe drinking water, sampling for *E. coli* can verify that controls are effective or identify fecal contamination risks that are as yet uncontrolled. Based on the analysis presented, we recommend the following priorities for sampling:

1. Prioritize sampling in rainy periods. If a system is otherwise deemed to be safe from fecal contamination based on low sanitary inspection hazards, sampling should be prioritized during rainy periods when the contamination risk is highest.
2. Develop sampling to progressively identify different contamination pathways. If a system is deemed safe from fecal contamination based on low sanitary inspection hazards and has no *E. coli* detected in the first sample, then plan subsequent samples to address different contamination pathways. We identified different pathways associated with rainfall—for example, the difference between rapid contamination pathways associated with heavy rainfall and contamination traveling further distances due to sustained rainfall. A third sample might address dry season risks, noting that contamination pathways will also differ depending on the antecedent dry period.

Sampling in rainy seasons can be difficult for logistical reasons in small systems, particularly in rural settings without good road access. Approaches that enable onsite testing, such as Aquagenx CBT, or the development of rural laboratories[45,46] support operational monitoring, reducing travel challenges that may be associated with using accredited laboratory facilities. If all sampling is focused on rainy seasons, the intensity of sampling and analysis activities may create a challenge for building and maintaining capacity. For these practical reasons, adding dry season samples to explore different risks may be appropriate rather than conducting further wet-season sampling.

Finally, the criteria for defining a PoC or water supply as safe should recognize the intermittent nature of fecal contamination and require preventative actions as well as at least seasonal testing for *E. coli*. In developing this work, it has become clear to us that the language for describing drinking water contamination is limited; and for that reason, we adopted the term "intermittent contamination", careful not to suggest common patterns in the name.

In summary, *E. coli* contamination in untreated water systems is intermittent, influenced by a variety of local and weather-related drivers but irregular in nature. Isolated sampling, including cross-sectional sampling campaigns—such as those that inform international monitoring—substantially underrepresents the presence of fecal contamination due to these irregular patterns. Improving water safety for the estimated 4.4 billion people who lack access to safe drinking water requires specific attention to small water systems, where the health burden is significant but largely unquantified due to its distributed nature. Implementing a risk-based approach to *E. coli* monitoring that employs cross-temporal repeated sampling, integrated with actions to improve water safety, will help prioritize resources to reduce fecal contamination in drinking water in small systems. Reframing *E. coli* monitoring as a risk-based approach within an operational water safety program will support progress toward safer drinking water.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsestwater.6c00247.

Supporting Information 1: sampling, analysis, and weather. Supporting Information 2: details of analysis methods. Supporting Information 3: analysis of frequency of sequences. Supporting Information 4: MICS data analysis (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Katrina J. Charles** − *School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, United Kingdom of Great Britain and Northern Ireland*; orcid.org/0000-0002-2236-7589; Email: Katrina.charles@ouce.ox.ac.uk

### Authors

**Raphaela Betz** − *School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, United Kingdom of Great Britain and Northern Ireland*

**Saskia Nowicki** − *School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, United Kingdom of Great Britain and Northern Ireland*; orcid.org/0000-0001-9011-6901

**Edith Darin** − *Leverhulme Centre for Demographic Science, Nuffield Department of Population Health, University of Oxford, Oxford OX1 2JD, United Kingdom of Great Britain and Northern Ireland*




https://doi.org/10.1021/acsestwater.6c00247
ACS EST Water XXXX, XXX, XXX−XXX

**Jamie Bartram** − *School of Civil Engineering, University of Leeds, Leeds LS2 9JT, United Kingdom of Great Britain and Northern Ireland; Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514, United States;* orcid.org/0000-0002-6542-6315

Complete contact information is available at: https://pubs.acs.org/10.1021/acsestwater.6c00247

## Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

## Notes

Thematic section: Occurrence, Fate, and Transport of Aquatic and Terrestrial Contaminants

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

K.J.C. and S.N. were supported by UK Aid from the UK Foreign, Commonwealth and Development Office (FCDO) (Programme Code 201880). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. E.D. was supported by the Leverhulme Trust under Grant RC-2018-003 for the Leverhulme Centre for Demographic Science. E.D. gratefully acknowledges the resources provided by the International Max Planck Research School for Population, Health and Data Science (IMPRS-PHDS).

## ■ REFERENCES

(1) WHO, UNICEF. *Progress on Household Drinking Water, Sanitation and Hygiene 2000–2024: special Focus on Inequalities*; Geneva, 2025. https://washdata.org/report/jmp-2025-wash-households. (5 Jan. 2026, date last accessed).

(2) Greenwood, E. E.; Lauber, T.; van den Hoogen, J.; Donmez, A.; Bain, R. E. S.; Johnston, R.; Crowther, T. W.; Julian, T. R. Mapping safe drinking water use in low- and middle-income countries. *Science* **2024**, *385* (6710), 784–790.

(3) Institute for Health Metrics and Evaluation (IHME). *GBD Compare Data Visualization*; Seattle: WA, 2025. https://www.healthdata.org/data-tools-practices/interactive-visuals/gbd-compare. (30 Jan. 2026, date last accessed).

(4) Prüss-Ustün, A.; Wolf, J.; Bartram, J.; et al. Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: An updated analysis with a focus on low- and middle-income countries. *Int. J. Hyg. Environ. Health* **2019**, *222* (5), 765.

(5) Hunter, P. R.; Zmirou-Navier, D.; Hartemann, P. Estimating the impact on health of poor reliability of drinking water interventions in developing countries. *Sci. Total Environ.* **2009**, *407*, 2621–2624.

(6) WHO. *Guidelines for Drinking-Water Quality. Fourth Edition Incorporating the First and Second Addenda.* World Health Organization, Geneva, 2022.

(7) Bain, R.; Johnston, R.; Khan, S.; et al. Monitoring Drinking Water Quality in Nationally Representative Household Surveys in Low- and Middle-Income Countries: Cross-Sectional Analysis of 27 Multiple Indicator Cluster Surveys 2014–2020. *Environ. Health Perspect.* **2021**, *129* (9), 097010.

(8) Charles, K. J.; Nowicki, S.; Bartram, J. K. A framework for monitoring the safety of water services: from measurements to security. *Npj Clean Water* **2020**, *3* (1), 36.

(9) O'Meara, W. P.; Collins, W. E.; McKenzie, F. E. Parasite Prevalence: A Static Measure of Dynamic Infections. *Am. J. Trop. Med. Hyg.* **2007**, *77* (2), 246.

(10) Hergott, D. E. B.; Owalla, T. J.; Staubus, W. J.; Seilie, A. M.; Chavtur, C.; Balkus, J. E.; Apio, B.; Lema, J.; Cemeri, B.; Akileng, A.; et al. Assessing the daily natural history of asymptomatic Plasmodium infections in adults and older children in Katakwi, Uganda: a longitudinal cohort study. *Lancet Microbe* **2024**, *5* (1), No. e72–80.

(11) Spencer, E.; Heneghan, C. *Prevalence-incidence (Neyman) bias. In: catalog of Bias*. 2017. https://catalogofbias.org/biases/prevalence-incidence-neyman-bias/. (29 May 2026, date last accessed).

(12) Kostyla, C.; Bain, R.; Cronk, R.; et al. Seasonal variation of fecal contamination in drinking water sources in developing countries: A systematic review. *Sci. Total Environ.* **2015**, *514*, 333–343.

(13) Ronoh, P.; Furlong, C.; Kansiime, F.; et al. Are There Seasonal Variations in Faecal Contamination of Exposure Pathways? An Assessment in a Low-Income Settlement in Uganda. *Int. J. Environ. Res. Public Health* **2020**, *17* (17), 6355.

(14) Ward, J. S. T.; Lapworth, D. J.; Read, D. S.; Pedley, S.; Banda, S. T.; Monjerezi, M.; Gwengweya, G.; MacDonald, A. M. Large-scale survey of seasonal drinking water quality in Malawi using in situ tryptophan-like fluorescence and conventional water quality indicators. *Sci. Total Environ.* **2020**, *744*, No. e140674.

(15) Dongzagla, A.; Jewitt, S.; O'Hara, S. Seasonality in faecal contamination of drinking water sources in the Jirapa and Kassena-Nankana Municipalities of Ghana. *Sci. Total Environ.* **2021**, *752*, 141846.

(16) Kumpel, E.; Cock-Esteb, A.; Duret, M.; et al. Seasonal variation in drinking and domestic water sources and quality in Port Harcourt, Nigeria. *Am. J. Trop. Med. Hyg.* **2017**, *96* (2), 437–445.

(17) Levy, K.; Hubbard, A. E.; Nelson, K. L.; et al. Drivers of Water Quality Variability in Northern Coastal Ecuador. *Environ. Sci. Technol.* **2009**, *43* (6), 1788–1797.

(18) Nguyen, K. H.; Operario, D. J.; Nyathi, M. E.; et al. Seasonality of drinking water sources and the impact of drinking water source on enteric infections among children in Limpopo, South Africa. *Int. J. Hyg. Environ. Health* **2021**, *231*, 113640.

(19) Charles, K. J.; Howard, G.; Villalobos Prats, E.; et al. Infrastructure alone cannot ensure resilience to weather events in drinking water supplies. *Sci. Total Environ.* **2022**, *813* (xxxx), 151876.

(20) Okotto-Okotto, J.; Wanza, P.; Kwoba, E.; Yu, W.; Dzodzomenyo, M.; Thumbi, S. M.; da Silva, D. G.; Wright, J. A.; et al. An Assessment of Inter-Observer Agreement in Water Source Classification and Sanitary Risk Observations. *Expo. Health* **2020**, *12*, 809–822.

(21) Setty, K. E.; Enault, J.; Loret, J. F.; et al. Time series study of weather, water quality, and acute gastroenteritis at Water Safety Plan implementation sites in France and Spain. *Int. J. Hyg. Environ. Health* **2018**, *221* (4), 714–726.

(22) Poulin, C.; Peletz, R.; Ercumen, A.; et al. What Environmental Factors Influence the Concentration of Fecal Indicator Bacteria in Groundwater? Insights from Explanatory Modeling in Uganda and Bangladesh. *Environ. Sci. Technol.* **2020**, *54* (21), 13566–13578.

(23) Powers, J. E.; Mureithi, M.; Mboya, J.; et al. Effects of High Temperature and Heavy Precipitation on Drinking Water Quality and Child Hand Contamination Levels in Rural Kenya. *Environ. Sci. Technol.* **2023**, *57* (17), 6975–6988.

(24) Rochelle-Newall, E.; Nguyen, T. M. H.; Le, T. P. Q.; Sengtaheuanghoung, O.; Ribolzi, O. A short review of fecal indicator bacteria in tropical aquatic ecosystems: Knowledge gaps and future directions. *Front. Microbiol.* **2015**, *6*, No. e308.

(25) Stukel, T.; Greenberg, E.; Dain, B.; Reed, F.; Jacobs, N. A Longitudinal Study of Rainfall and Coliform Contamination in Small Community Drinking Water Supplies. *Environ. Sci. Technol.* **1990**, *24*, 571–575.

(26) Hubbard, S.; Wolf, J.; Oza, H. H.; Arnold, B. F.; Freeman, M. C.; Levy, K. Differential Effectiveness of Water, Sanitation, and Handwashing Interventions to Reduce Child Diarrhea in Dry and Rainy Seasons: A Systematic Review and Meta-Analysis of Intervention Trials. *Environ. Health Perspect.* **2025**, *133* (2), 26001.

(27) Howard, G.; Pedley, S.; Barrett, M.; et al. Risk factors contributing to microbiological contamination of shallow groundwater in Kampala, Uganda. *Water Res.* **2003**, *37* (14), 3421−3429.
(28) Wu, J.; Yunus, M.; Islam, M. S.; et al. Influence of Climate Extremes and Land Use on Fecal Contamination of Shallow Tubewells in Bangladesh. *Environ. Sci. Technol.* **2016**, *50* (5), 2669−2676.
(29) O'Dwyer, J.; Hynds, P. D.; Byrne, K. A.; et al. Development of a hierarchical model for predicting microbiological contamination of private groundwater supplies in a geologically heterogeneous region. *Environ. Pollut.* **2018**, *237*, 329−338.
(30) Nowicki, S. *Data, decisions, and drinking water safety: an interdisciplinary analysis of the complex adaptive response to monitoring in rural Kenya*; PhD thesis, University of Oxford; 2021. https://ora.ox.ac.uk/objects/uuid:952a1144-59e7-4cc7-8bb9-4273556c3892.
(31) Nowicki, S.; Lapworth, D. J.; Ward, J. S. T.; et al. Tryptophan-like fluorescence as a measure of microbial contamination risk in groundwater. *Sci. Total Environ.* **2019**, *646*, 782−791.
(32) WHO. *Guidelines for Drinking-Water Quality: small Water Supplies*; World Health Organization: Geneva, 2024. https://iris.who.int/server/api/core/bitstreams/4a3dc8a4-9f08-4d62-b7ec-96d70be6b177/content. (15 Feb. 2024, date last accessed).
(33) European Union *Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption, L 435/1*; 2020. https://eur-lex.europa.eu/eli/dir/2020/2184/oj.
(34) Wagner, J.; Merner, S.; Innocenti, S.; et al. Can solar water kiosks generate sustainable revenue streams for rural water services? *World Dev.* **2025**, *185*, 106787.
(35) Trimmer, J. T.; Delaire, C.; Marshall, K.; et al. Centralized or Onsite Testing? Examining the Costs of Water Quality Monitoring in Rural Africa. *Environ. Sci. Technol.* **2024**, *58* (26), 11236−11246.
(36) Charles, K. J.; Nowicki, S.; Ong, L. A.; et al. Interpreting drinking water quality samples: Understanding contamination pathways at the point of collection. *PLOS Water* **2025**, *4* (10), No. e0000373.
(37) Liao, T. F.; Bolano, D.; Brzinsky-Fay, C.; et al. Sequence analysis: Its past, present, and future. *Soc. Sci. Res.* **2022**, *107*, 102772.
(38) Fu, L.; Wang, Y. G. Statistical Tools for Analyzing Water Quality Data. In *Water Quality Monitoring and Assessment*; InTech, 2012. DOI: .
(39) Lorimer, M. F.; Kiermeier, A. Analysing microbiological data: Tobit or not Tobit? *Int. J. Food Microbiol.* **2007**, *116* (3), 313−318.
(40) Gabadinho, A.; Ritschard, G.; Müller, N. S.; et al. Analyzing and visualizing state sequences in R with TraMineR. *J. Stat. Software* **2011**, *40* (4), 1−37.
(41) Ritschard, G. Measuring the Nature of Individual Sequences, article. *Sociol Methods Res.* **2023**, *52* (4), 2016−2049.
(42) Central Bureau of Statistics (CBS). *Nepal Multiple Indicator Cluster Survey 2019, Survey Findings Report*; Kathmandu, Nepal, 2020. https://washdata.org/report/nepal-2019-mics-0. (5 Jan. 2026, date last accessed).
(43) Bangladesh Bureau of Statistics, UNICEF Bangladesh. *Bangladesh MICS 2019 Water Quality Thematic Report*; Dhaka, Bangladesh, 2021. https://psb.gov.bd/policies/wqtr2019.pdf.
(44) Charles, K.; Nowicki, S.; Armstrong, A.; Hope, R.; McNicholl, D.; Nilsson, K. Results-Based Funding for Safe Drinking Water Services: How a Standard Contract Design with Payment for Results Can Accelerate Safe Drinking Water Services at Scale. In *REACH Working Paper*; University of Oxford and Uptime Global: Oxford, UK, 2023.
(45) Nowicki, S.; Muturi, J.; Boller, M.; Uprety, S.; Kunwar, B. M.; Nyaga, C.; Sammy, M.; Rahaman, M. F.; Osman, N.; Marks, S. J.et al.Enabling data for decision making on drinking water safety through fit-for-purpose laboratories*Environ. Sci.: Water Res. Technol.***2026**
(46) Wright, J.; Liu, J.; Bain, R.; et al. Water quality laboratories in Colombia: A GIS-based study of urban and rural accessibility. *Sci. Total Environ.* **2014**, *485−486*, 643−652.