Can Sanitary Inspection Surveys Predict Risk of Microbiological Contamination of Groundwater Sources? Evidence from Shallow Tubewells in Rural Bangladesh

Ayşe Ercumen,^{1*} Abu Mohd Naser,^{2,3} Benjamin F. Arnold,¹ Leanne Unicomb,² John M. Colford Jr.,¹ and Stephen P. Luby^{2,4,5}

¹Division of Epidemiology, School of Public Health, University of California, Berkeley, California; ²International Centre for Diarrhoeal Disease Research, Bangladesh (icddr), Dhaka, Bangladesh; ³Rollins School of Public Health, Emory University, Atlanta, Georgia;

⁴School of Medicine, Stanford University, Stanford, California; ⁵Centers for Disease Control and Prevention, Atlanta, Georgia

Abstract. Accurately assessing the microbiological safety of water sources is essential to reduce waterborne fecal exposures and track progress toward global targets of safe water access. Sanitary inspections are a recommended tool to assess water safety. We collected 1,684 water samples from 902 shallow tubewells in rural Bangladesh and conducted sanitary surveys to assess whether sanitary risk scores could predict water quality, as measured by *Escherichia coli*. We detected *E. coli* in 41% of tubewells, mostly at low concentrations. Based on sanitary scores, 31% of wells were low risk, 45% medium risk, and 25% high or very high risk. Older wells had higher risk scores. *Escherichia coli* levels were higher in wells where the platform was cracked or broken ($\Delta\log_{10} = 0.09, 0.00\text{--}0.18$) or undercut by erosion ($\Delta\log_{10} = 0.13, 0.01\text{--}0.24$). However, the positive predictive value of these risk factors for *E. coli* presence was low ($< 50\%$). Latrine presence within 10 m was not associated with water quality during the wet season but was associated with less frequent *E. coli* detection during the dry season (relative risk = 0.72, 0.59–0.88). Sanitary scores were not associated with *E. coli* presence or concentration. These findings indicate that observed characteristics of a tubewell, as measured by sanitary inspections in their current form, do not sufficiently characterize microbiological water quality, as measured by *E. coli*. Assessments of local groundwater and geological conditions and improved water quality indicators may reveal more clear relationships. Our findings also suggest that the dominant contamination route for shallow groundwater sources is short-circuiting at the wellhead rather than subsurface transport.

INTRODUCTION

The Millennium Development Goals targeted to halve the proportion of the world population without access to safe drinking water by 2015. This goal was met in 2010, and the post-2015 Sustainable Development Goals (SDGs) target “universal and equitable access to safe and affordable drinking water for all.”¹ Classifying water sources as “improved” versus “unimproved” is a commonly used yardstick for tracking progress toward these targets. However, this definition is based on source type and does not adequately capture whether water from a source is safe to drink. Although improved sources typically have better water quality than unimproved sources, they can still contain fecal contamination.² The Joint Monitoring Programme for Water and Sanitation reports that 663 million people worldwide use an unimproved water source as of 2015.¹ Yet, studies that account for water quality and sanitary risks estimate that 1.8 billion people obtain water from fecally contaminated sources,^{2,3} and the SDGs place emphasis on “safely managed” drinking water.⁴ Standardized sanitary inspections have been developed as a complementary tool to assess the safety of water sources, and their use has been recommended to augment the definition of safe water along with source type and water quality measurements. However, a recent review found no association between sanitary scores and water quality as measured by fecal indicators.²

Tubewells (boreholes) drawing groundwater constitute an improved source and serve an estimated 17% of the world’s population.⁵ While typically safer than surface water, recent

reviews suggest that 43% of tubewells show evidence of fecal contamination, sometimes at high levels.^{2,6} Sanitary inspection scores have been developed for tubewells to capture potential contamination pathways, including the presence of latrines in the vicinity of the well and the physical integrity of various wellhead components such as the drainage channel, concrete platform, handpump, and trunk.⁷ However, there is mixed evidence on how well sanitary inspections predict risk of fecal contamination of shallow groundwater sources.^{7–12}

In Bangladesh, shallow tubewells are the primary source of drinking water and contain fecal indicator bacteria as well as human enteric pathogens.^{13–18} Mechanisms that lead to this contamination include infiltration from ponds and latrines,^{19,20} short-circuiting at the wellhead through unsealed parts,²¹ and biofilm formation on the handpump.²² A previous study in rural Bangladesh found no correlation between sanitary scores that quantify these pathways and tubewell water quality.⁷ However, this study was based on a one-time water sample collected during the rainy season. Tubewell water quality varies significantly with rainfall in Bangladesh, with higher levels of contamination during the rainy season.^{18,23} A one-time sample collected during a single season may therefore be an inadequate indicator given the temporal variation in tubewell water quality. The association between sanitary risk factors and tubewell water quality may also be modified by seasonal factors such as the elevation of the groundwater table. Additionally, climate change is expected to alter weather patterns globally; understanding how water sources are affected by sanitary risk factors under different seasonal conditions is important to assess and mitigate the impact of climate change on drinking water safety.

Herein, we use data from 902 tubewells in rural Bangladesh sampled once during the dry season and once during the

*Address correspondence to Ayşe Ercumen, Division of Epidemiology, School of Public Health, University of California, Berkeley, 50 University Hall, No. 7360, Berkeley, CA 94720. E-mail: aercumen@berkeley.edu

rainy season to assess whether sanitary risk scores are associated with microbiological contamination of tubewell water and whether dominant sanitary risk factors vary by season.

MATERIALS AND METHODS

Data collection. The study was conducted in 87 villages in the Mymensingh District of central Bangladesh, among households that consistently relied on a shallow tubewell (< 250 ft) as their primary source of drinking water. We collected 100 mL tubewell water samples over the course of 1 year (October 2011–November 2012) as part of a randomized controlled trial of water treatment in rural Bangladesh, analyzed them for *Escherichia coli* using the U.S. Environmental Protection Agency Method 1604, and expressed *E. coli* counts in colony-forming units (CFU) per 100 mL.²⁴ We analyzed 10% blanks and 10% duplicates for quality control. Further details of the trial and the water quality analyses are reported elsewhere.¹⁸

We aimed to sample approximately 900 wells during the dry season (November–May) and resample the same wells during the wet season (June–October) to assess seasonal variation. In central Bangladesh, it is consistently rainy during the monsoon season and consistently dry during the rest of the year, with > 80% of the annual rainfall occurring during the wet season.²⁵ Other work in the study area demonstrated that using calendar months to define wet versus dry seasons adequately captures recent rainfall; during days classified as “wet season” based on month, 91% of observations had rain within the last 2 days, and during days classified as “dry season,” 96% of observations had no rain within the last 2 days (A. Ercumen, unpublished data).

As part of the baseline data collection for the trial, field research assistants conducted a sanitary inspection for each tubewell, following classroom and field training conducted by the study investigators on identifying sanitary risk factors. We calculated a composite sanitary risk score based on these observations (Table 1), and classified wells with a sanitary score ≤ 3 as low risk, 4–5 as medium risk, 6–7 as high risk, and 8–10 as very high risk, consistent with previous uses of the tubewell sanitary risk score.⁷ We also collected information on reported tubewell age and depth, and household sanitation infrastructure and practices, including observed presence of an improved latrine, and reported location of defecation and feces disposal for children < 2 years of age.

Statistical methods. We calculated the percentage of *E. coli*-positive samples, samples with *E. coli* exceeding 10 CFU/100 mL, and the $\log_{10}$ *E. coli* concentration, sub-

stituting a value of 0.5 for samples with < 1 CFU/100 mL to calculate the logarithm. Samples with > 200 colonies were classified as “too numerous to count”; we substituted a value of 200 CFU/100 mL for these in the statistical analyses. We defined the wet season as June–October and the dry season as November–May, and compared the percentage of samples with *E. coli* using McNemar’s χ^2 tests and $\log_{10}$ *E. coli* concentrations using paired *t* tests between the wet and dry seasons. We assessed the relationship between sanitary risk factors and *E. coli* presence (≥ 1 CFU/100 mL), *E. coli* exceeding the World Health Organization (WHO) low-risk threshold (> 10 CFU/100 mL), and $\log_{10}$ *E. coli* concentration using generalized linear models with robust standard errors to account for multiple samples collected from the same well, using a log link and binomial distribution for *E. coli* presence and an identity link and normal distribution for $\log_{10}$ concentration.²⁶ We examined the impact of individual components of the sanitary score as well as the composite score on water quality. For risk factors associated with detection of *E. coli* at the $P < 0.10$ level, we calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the risk factor as a predictor of *E. coli* presence. We assessed the effect of household sanitary conditions, such as access to improved sanitation and safe child feces handling practices, and additional tubewell characteristics, such as well depth and age, on tubewell water quality. We also assessed whether tubewell age is correlated with sanitary risk factors by using *t* tests and analysis of variance (ANOVA) to compare mean well age across physical well characteristics and sanitary risk categories. We explored effect modification by season by including an interaction term for wet versus dry season in the regressions.

RESULTS

Tubewell water quality. We successfully collected two samples from 756 wells, whereas 133 wells were sampled only once and 13 wells were sampled three times, leading to a total of 1,684 water samples from 902 unique wells. Of the 756 wells sampled twice, 718 were successfully sampled once per dry versus rainy season; the remaining 38 were sampled twice in the same season. Quality control objectives for absence of background contamination (blanks) and precision (duplicates) were met. Of the 1,684 samples, 41% were positive for *E. coli* (≥ 1 CFU/100 mL) and 14% exceeded the WHO low-risk threshold of 10 CFU/100 mL (Table 2).²⁷ A small number (2%) of the samples

TABLE 1
Tubewell sanitary inspection score

Risk factor	Score (yes = 1, no = 0)
1. Is there a latrine within 10 m of the tubewell?	Yes/No
2. Are there any other sources of pollution within 10 m of the tubewell (e.g., cow sheds, fertilizers)?	Yes/No
3. Is there a ditch or pond within 10 m of the tubewell?	Yes/No
4. Is the drainage faulty allowing ponding within 2 m of the tubewell?	Yes/No
5. Is the drainage channel cracked, broken, or in need of cleaning (i.e., large pieces of debris blocking the channel)?	Yes/No
6. Is the platform smaller than 5 × 5 ft in dimension?	Yes/No
7. Is the platform cracked, broken, or in need of cleaning?	Yes/No
8. Is the platform undercut by erosion?	Yes/No
9. Is the handpump loose at the point of attachment?	Yes/No
10. Is the trunk loose at the point of attachment?	Yes/No

TABLE 2
Escherichia coli in tubewell samples by season

	N	≥ 1 CFU	> 10 CFU	Log ₁₀ <i>E. coli</i>
		n (%)	n (%)	Mean (SD)
Year-round	1,684	689 (41)	235 (14)	0.14 (0.70)
Wet season (June–October)	810	398 (49)	147 (19)	0.25 (0.76)
Dry season (November–May)	874	291 (33)	88 (10)	0.03 (0.62)

CFU = colony-forming units; SD = standard deviation.

were over the detection limit of 200 CFU/100 mL. During the wet season, 49% of wells were *E. coli* positive compared with 33% during the dry season ($P < 0.005$, Table 2). The highest percentages of positive samples were observed during the months of June–September (Figure 1). Log₁₀ *E. coli* concentrations showed a similar trend with higher counts in the wet season ($P < 0.005$, Table 2). Of the 718 wells sampled once per season, 238 (33%) had no detectable *E. coli* in either season, 117 (16%) had detectable *E. coli* in both seasons, 242 (34%) had detectable *E. coli* during the wet season only, and 121 (17%) during the dry season only. There was no correlation between *E. coli* detection among multiple samples from the same well (intraclass correlation coefficient = 0.00).

Tubewell sanitary risk score. Approximately half of the tubewells were located within 10 m of a latrine or pond (Table 3); 83% of tubewells were located in a compound that had a latrine, and 35% in a compound that had an improved latrine. Among compounds with tubewells tested, safe child defecation and feces disposal practices were infrequently reported by household members. Approximately a quarter of the tubewells lacked a drainage channel and half had a channel that was broken or in need of cleaning. Thirty-six percent of tubewells lacked a concrete platform, 48% had a small ($< 5 \times 5$ ft) platform, 25% had a cracked or broken platform, and 11% had a platform that was undercut by erosion. Thirty-two percent of the wells had a loose handpump and 6% had a loose trunk. Reported well depth ranged from 22 to 250 ft and well age ranged from < 1 to 30 years (Table 3). Tubewell sanitary scores were evenly spread. The median score was 4 (range: 0–9), and 31% of

wells were low risk, 45% medium risk, 22% high risk, and only 3% very high risk (Table 3).

Higher tubewell age was correlated with deteriorated physical wellhead characteristics. Older wells were more likely to have a drainage channel or platform, but these were more likely to be cracked, broken, or eroded, and the handpumps more likely to be loose (all t test P values < 0.005 , Table 4). Older wells were less likely to be located within 10 m of a latrine (t test P value < 0.005 , Table 4). Increasing well age was correlated with higher sanitary risk; the mean age was 7.2 years for low-risk wells, 7.7 years for medium-risk wells, 9.3 years for high-risk wells, and 10.6 years for very high-risk wells (ANOVA P value < 0.005).

Sanitary risk score and water quality. Faulty wellhead components were associated with compromised tubewell water quality; wells with cracked, broken, or eroded platforms and loose handpumps were 12–17% more likely to contain *E. coli*, though the associations were borderline nonsignificant (Table 5). *Escherichia coli* concentrations were higher in wells where the platform was cracked or

TABLE 3
Sanitary score, tubewell characteristics, and household sanitary conditions

Sanitary score components	%	n (range)	N
Latrine within 10 m	42	374	899
Other source of pollution within 10 m	50	447	899
Pond within 10 m	58	520	899
Poor drainage/pooling within 2 m	58	518	897
No drainage channel	25	225	898
Dirty or broken drainage channel	47	425	898
No platform	36	322	897
Small ($< 5 \times 5$ ft) platform	48	429	897
Cracked or broken platform	25	220	897
Eroded platform	11	100	897
Loose handpump	32	265	832
Loose trunk	6	53	844
Composite sanitary score			
Median score	4	(0–9)	830
Low risk (0–3)	31	254	830
Medium risk (4–5)	45	373	830
High risk (6–7)	22	182	830
Very high risk (8–10)	3	21	830
Tubewell characteristics			
Mean tubewell depth (ft)	147	(22–250)	879
Mean tubewell age (years)	8.3	(< 1–30)	897
Priming	9	85	902
Household sanitary conditions			
Latrine in compound	83	750	902
Improved latrine in compound	35	312	902
Children < 2 years defecate in latrine or potty	5	44	902
Feces of children < 2 years disposed of in latrine	12	112	901

FIGURE 1. Percent of *Escherichia coli*-positive tubewell samples by month.

TABLE 4
Average tubewell age by sanitary score characteristics

	Tubewell age (years)		t test P value
	Yes	No	
Latrine within 10 m	6.8	9.3	< 0.005
Other source of pollution within 10 m	8.0	8.5	0.24
Pond within 10 m	8.2	8.4	0.67
Poor drainage/pooling within 2 m	8.7	7.7	0.05
Lacking drainage channel	7.0	8.7	< 0.005
Dirty or broken drainage channel	9.5	7.4	< 0.005
Lacking platform	6.7	9.1	< 0.005
Small (< 5 × 5 ft) platform	8.3	11.6	< 0.005
Cracked or broken platform	12.4	7.1	< 0.005
Eroded platform	13.2	8.3	< 0.005
Loose handpump	10.1	6.9	< 0.005
Loose trunk	7.1	8.0	0.34

broken ($\Delta\log_{10} = 0.09$, 0.00–0.18, $P = 0.04$) or undercut by erosion ($\Delta\log_{10} = 0.13$, 0.01–0.24, $P = 0.03$, Table 6). Increasing tubewell age significantly increased the probability of a well containing *E. coli* as well as the *E. coli* concentration ($\Delta\log_{10} = 0.07$, 0.03–0.12, for every 10-year increase in well age, $P < 0.005$, Table 6); well age seemed to be a proxy for physical integrity of wellhead components, with older wells more likely to have faulty drainage channels or platforms (Table 4). The sensitivity of observed wellhead characteristics against measured *E. coli* detection ranged from 20% to 41%, specificity from 57% to 84%, PPV from 38% to 48%, and NPV from 57% to 62% (Table 7).

Surprisingly, tubewells that had a latrine within 10 m were less likely to be positive for *E. coli* (relative risk [RR] = 0.88, 0.79–0.99, $P = 0.04$, Table 5). A similar but nonsignificant association remained once we controlled for well age (adjusted RR = 0.90, 0.80–1.02, $P = 0.09$, Supplemental Tables 1 and 2). Neither presence of a latrine or an improved latrine in the compound nor child defecation and feces disposal practices were associated with tubewell water quality (Tables 5 and 6). The composite sanitary score, both in categorical form and in binary comparison of high/very high- versus low-risk wells and medium- versus low-risk wells, was unassociated with *E. coli* presence or concentration (Tables 5 and 6). Using *E. coli* exceeding 10 CFU/100 mL as the outcome measure yielded similar results (Supplemental Table 3).

Subgroup analysis by season showed that tubewells with a latrine within 10 m were less likely to be *E. coli* positive in the dry season (RR = 0.72, 0.59–0.88, Table 5), whereas there was no association between latrine presence and *E. coli* detection in the wet season (interaction P value < 0.005, Table 5). Controlling for tubewell age did not change these associations (Supplemental Table 1). An eroded platform was a risk factor only in the dry season (RR = 1.41, 1.08–1.84, interaction P value = 0.06, Table 5), whereas tubewell age was a risk factor only in the wet season ($\Delta\log_{10} = 0.14$, 0.06–0.21, for every 10-year increase in well age, interaction P value = 0.01, Table 6). Surprisingly, lack of a drainage channel, lack of a platform, and a loose trunk were associated with significantly lower *E. coli* counts

TABLE 5
Detection of *Escherichia coli* (≥ 1 CFU) vs. sanitary risk score, tubewell characteristics, and household sanitary conditions

<i>E. coli</i> ≥ 1 CFU	Both seasons		Dry season		Wet season		Interaction
	RR (95% CI)	<i>P</i> value	RR (95% CI)	<i>P</i> value	RR (95% CI)	<i>P</i> value	<i>P</i> value
Sanitary score							
Latrine within 10 m	0.88 (0.79, 0.99)	0.04**	0.72 (0.59, 0.88)	0.00**	1.04 (0.90, 1.20)	0.62	0.00
Other source of pollution within 10 m	1.02 (0.91, 1.14)	0.78	1.01 (0.83, 1.22)	0.94	1.02 (0.88, 1.17)	0.81	0.93
Pond within 10 m	1.09 (0.97, 1.23)	0.14	1.07 (0.88, 1.30)	0.50	1.11 (0.96, 1.28)	0.17	0.77
Poor drainage/pooling within 2 m	1.04 (0.92, 1.16)	0.54	1.07 (0.88, 1.30)	0.48	1.01 (0.87, 1.16)	0.93	0.62
No drainage channel (vs. present)	0.96 (0.85, 1.10)	0.58	1.05 (0.84, 1.30)	0.67	0.90 (0.76, 1.06)	0.21	0.29
Drainage channel broken or dirty (vs. intact)	1.00 (0.87, 1.14)	0.95	1.05 (0.83, 1.33)	0.66	0.96 (0.81, 1.13)	0.62	0.51
No platform (vs. present)	0.92 (0.81, 1.04)	0.16	0.98 (0.80, 1.20)	0.86	0.87 (0.75, 1.02)	0.08*	0.35
Platform smaller than 5 × 5 ft (vs. larger than 5 × 5 ft)	0.97 (0.82, 1.14)	0.72	0.83 (0.64, 1.08)	0.17	1.11 (0.91, 1.36)	0.30	0.08
Platform cracked or broken (vs. intact)	1.14 (0.99, 1.32)	0.06*	1.11 (0.88, 1.42)	0.38	1.16 (0.98, 1.37)	0.09*	0.80
Platform eroded (vs. not eroded)	1.17 (0.99, 1.38)	0.06*	1.41 (1.08, 1.84)	0.01**	1.01 (0.80, 1.26)	0.96	0.06
Handpump loose	1.12 (0.99, 1.27)	0.08*	1.12 (0.91, 1.38)	0.29	1.12 (0.96, 1.32)	0.15	0.98
Trunk loose	0.99 (0.78, 1.26)	0.94	1.17 (0.81, 1.70)	0.41	0.86 (0.61, 1.22)	0.40	0.27
Sanitary score	1.03 (0.99, 1.07)	0.20	1.02 (0.95, 1.09)	0.55	1.03 (0.98, 1.08)	0.22	0.83
Medium sanitary score (vs. low)	1.03 (0.89, 1.20)	0.68	0.96 (0.75, 1.23)	0.75	1.09 (0.91, 1.31)	0.36	0.42
High/very high sanitary score (vs. low)	1.14 (0.97, 1.34)	0.12	1.10 (0.84, 1.43)	0.49	1.17 (0.95, 1.44)	0.13	0.70
Other tubewell characteristics							
Tubewell depth (in increments of 10 ft)	1.00 (0.99, 1.02)	0.50	1.01 (0.99, 1.03)	0.48	1.00 (0.99, 1.02)	0.87	0.64
Tubewell age (in increments of 10 years)	1.10 (1.02, 1.18)	0.01**	1.01 (0.88, 1.15)	0.93	1.16 (1.07, 1.26)	0.00**	0.08
Priming	1.13 (0.95, 1.35)	0.18	1.26 (0.95, 1.68)	0.12	1.03 (0.82, 1.31)	0.78	0.30
Household sanitary conditions							
Latrine in compound (vs. no latrine)	0.95 (0.81, 1.11)	0.49	0.85 (0.67, 1.08)	0.18	1.02 (0.84, 1.24)	0.81	0.22
Improved latrine in compound (vs. unimproved or no latrine)	0.97 (0.86, 1.09)	0.58	0.92 (0.75, 1.13)	0.44	1.00 (0.86, 1.16)	0.99	0.54
Children < 2 years defecate in latrine or potty	0.92 (0.68, 1.26)	0.62	0.90 (0.57, 1.44)	0.67	0.96 (0.67, 1.38)	0.82	0.83
Feces of children < 2 years disposed of in latrine	0.99 (0.83, 1.18)	0.91	0.95 (0.71, 1.27)	0.72	1.02 (0.82, 1.27)	0.86	0.69

CFU = colony-forming units; CI = confidence interval; RR = relative risk.

* P value ≤ 0.10 .

** P value ≤ 0.05 .

TABLE 6
Log₁₀ *Escherichia coli* concentration vs. sanitary risk score, tubewell characteristics, and household sanitary conditions

Log ₁₀ <i>E. coli</i>	Both seasons		Dry season		Wet season		Interaction
	$\Delta\log_{10}$ (95% CI)	P value	$\Delta\log_{10}$ (95% CI)	P value	$\Delta\log_{10}$ (95% CI)	P value	P value
Sanitary score							
Latrine within 10 m	-0.05 (-0.11, 0.02)	0.16	-0.07 (-0.15, 0.01)	0.10*	-0.01 (-0.12, 0.10)	0.83	0.40
Other source of pollution within 10 m	0.03 (-0.03, 0.10)	0.32	0.03 (-0.05, 0.11)	0.46	0.03 (-0.08, 0.14)	0.56	0.99
Pond within 10 m	0.02 (-0.05, 0.09)	0.53	-0.01 (-0.10, 0.07)	0.78	0.06 (-0.05, 0.16)	0.30	0.33
Poor drainage/pooling within 2 m	0.02 (-0.05, 0.08)	0.62	0.02 (-0.06, 0.10)	0.64	0.01 (-0.10, 0.12)	0.85	0.89
No drainage channel (vs. present)	-0.03 (-0.10, 0.04)	0.42	0.05 (-0.05, 0.15)	0.35	-0.12 (-0.23, 0.00)	0.04**	0.04
Drainage channel broken or dirty (vs. intact)	0.00 (-0.08, 0.08)	1.00	0.07 (-0.03, 0.16)	0.18	-0.07 (-0.21, 0.06)	0.31	0.10
No platform (vs. present)	-0.04 (-0.11, 0.02)	0.21	0.03 (-0.06, 0.12)	0.52	-0.12 (-0.23, -0.01)	0.03**	0.04
Platform smaller than 5 × 5 ft (vs. larger than 5 × 5 ft)	-0.05 (-0.15, 0.05)	0.35	-0.09 (-0.21, 0.03)	0.14	0.02 (-0.14, 0.18)	0.81	0.28
Platform cracked or broken (vs. intact)	0.09 (0.00, 0.18)	0.04**	0.06 (-0.05, 0.17)	0.28	0.12 (-0.02, 0.26)	0.09*	0.48
Platform eroded (vs. not eroded)	0.13 (0.01, 0.24)	0.03**	0.16 (0.01, 0.32)	0.04**	0.08 (-0.11, 0.26)	0.42	0.49
Handpump loose	0.05 (-0.02, 0.13)	0.14	0.03 (-0.07, 0.12)	0.59	0.09 (-0.03, 0.21)	0.15	0.43
Trunk loose	-0.07 (-0.19, 0.05)	0.23	0.07 (-0.13, 0.26)	0.51	-0.22 (-0.40, -0.05)	0.01**	0.05
Sanitary score	0.01 (-0.01, 0.04)	0.26	0.01 (-0.02, 0.05)	0.41	0.01 (-0.02, 0.05)	0.48	0.97
Medium sanitary score (vs. low)	-0.01 (-0.09, 0.07)	0.85	-0.03 (-0.13, 0.07)	0.50	0.02 (-0.10, 0.15)	0.72	0.49
High/very high sanitary score (vs. low)	0.07 (-0.03, 0.16)	0.17	0.05 (-0.07, 0.18)	0.41	0.08 (-0.07, 0.24)	0.28	0.77
Other tubewell characteristics							
Tubewell depth (in increments of 10 ft)	0.00 (-0.01, 0.01)	0.91	0.00 (0.00, 0.01)	0.31	0.00 (-0.02, 0.01)	0.47	0.23
Tubewell age (in increments of 10 years)	0.07 (0.03, 0.12)	0.00**	0.01 (-0.05, 0.07)	0.71	0.14 (0.06, 0.21)	0.00**	0.01
Priming	0.00 (-0.10, 0.11)	0.94	-0.01 (-0.13, 0.11)	0.87	0.01 (-0.17, 0.19)	0.89	0.85
Household sanitary conditions							
Latrine in compound (vs. no latrine)	-0.01 (-0.10, 0.08)	0.79	-0.03 (-0.14, 0.08)	0.58	0.01 (-0.13, 0.15)	0.92	0.67
Improved latrine in compound (vs. unimproved or no latrine)	-0.03 (-0.10, 0.04)	0.43	-0.05 (-0.13, 0.04)	0.27	-0.01 (-0.12, 0.11)	0.92	0.57
Children < 2 years defecate in latrine or potty	-0.06 (-0.21, 0.08)	0.38	0.02 (-0.18, 0.21)	0.87	-0.15 (-0.33, 0.04)	0.12	0.20
Feces of children < 2 years disposed of in latrine	0.03 (-0.08, 0.14)	0.56	0.02 (-0.11, 0.15)	0.74	0.05 (-0.13, 0.22)	0.60	0.82

CI = confidence interval.

*P value ≤ 0.10.

**P value ≤ 0.05.

in the wet season but unassociated with *E. coli* concentration in the dry season (all interaction *P* values ≤ 0.05, Table 6).

DISCUSSION

Summary of findings. Using a large dataset of 1,684 water samples collected from 902 tubewells during the dry and rainy seasons in rural Bangladesh, we found that composite sanitary scores did not predict shallow groundwater

quality in this setting. These findings are consistent with previous studies that found no relationship between overall sanitary inspection scores and tubewells drawing shallow groundwater.^{7,8} Similar evidence also exists for other shallow groundwater sources. A study that sampled 25 protected springs monthly over a year in urban Uganda showed no correlation between sanitary risk scores and fecal indicator bacteria when population density was accounted for.⁹ Similarly, longitudinal sampling of 61 hand-dug wells over 6 years in peri-urban Kenya found no association between thermotolerant coliforms and sanitary scores.¹²

However, our findings demonstrate that, while the composite sanitary score was not associated with water quality, individual components of the risk score that measure the integrity of wellhead components were associated with the presence and concentration of *E. coli* in tubewells (Tables 5 and 6). This is in contrast to the previous assessment of tubewell sanitary inspections in Bangladesh, which cross-sectionally sampled 203 wells during the late wet season and found no association between water quality and individual risk factors included in the sanitary score.⁷ It is possible that our study's larger sample size and our sampling scheme that covered both the dry and rainy seasons allowed additional statistical power to detect these associations and assess effect modification by season. Indeed, the associations between wellhead characteristics and water quality varied with season in our study. Ultimately, however, the predictive power of these observed characteristics in

TABLE 7

Sensitivity, specificity, PPV, and NPV of sanitary risk factors for the presence of *Escherichia coli*

	Risk factor present (n)	Risk factor absent (n)	Sens (%)	Spec (%)	PPV (%)	NPV (%)
Latrine within 10 m						
<i>E. coli</i>	263	421	38	57	38	57
No <i>E. coli</i>	432	562				
Cracked or broken platform						
<i>E. coli</i>	185	261	41	64	45	60
No <i>E. coli</i>	222	396				
Eroded platform						
<i>E. coli</i>	90	356	20	84	48	59
No <i>E. coli</i>	99	519				
Loose handpump						
<i>E. coli</i>	208	406	34	70	43	62
No <i>E. coli</i>	278	656				

NPV = negative predictive value; PPV = positive predictive value.

distinguishing wells with compromised water quality was low, with PPVs and NPVs similar to what would be expected by chance.

Other evidence supports the importance of the structural integrity of the wellhead as a risk factor for the quality of groundwater sources. A study that sampled 233 wells in Mozambique showed that the quality of the wellhead components was more strongly associated with detection of thermotolerant coliforms in well water than the presence of a latrine within 30 m.¹⁰ Similarly, monthly monitoring of 25 wells over 12 months in Mozambique identified faulty wellhead components as a more significant risk factor than nearby latrines.¹¹

Our findings suggest that the presence of a latrine within a 10-m radius of the well was associated with less frequent detection of *E. coli* during the dry season (Tables 5 and 6). This could be because, in the absence of a latrine, open defecation likely leads to accumulation of fecal matter near the well during the dry season when there are no rains to regularly wash it away. Notably, in rural Bangladesh, open defecation by young children occurs even in the presence of a latrine (M. S. Islam, unpublished data) but likely presents a smaller fecal load than open defecation by multiple compound residents when there is no latrine. The absence of a latrine in the immediate vicinity of the tubewell could also mean that child and animal feces that accumulate near the well are left on the ground or swept aside instead of being disposed of in the latrine. Contaminated puddles in the wellhead catchment could then short-circuit into the well through faulty wellhead components. Such short-circuiting seems a more likely route of contamination than infiltration through the subsurface, given that there was no association between latrines and water quality in the wet season, during which the elevated groundwater table would provide ideal hydraulic conditions for subsurface transport of pathogens. Similarly, a study in Mozambique concluded that introduction of fecal contamination at the wellhead from child and animal feces present in the wellhead catchment was the dominant route of contamination for wells, rather than infiltration from latrines into groundwater aquifers.¹¹

There are a few possible explanations for why the presence of a latrine within 10 m of a tubewell or in the compound was not associated with increased contamination of tubewells during the wet season. Pathogens can travel distances on the order of 50 m in the subsurface depending on grain size, saturation of soil, and the type of pathogen.^{28,29} Even in the absence of a latrine within 10 m of a well, infiltration from latrines outside this radius can therefore still pose a threat. A study in Uganda showed that presence of a latrine within 30 m uphill of protected springs was significantly associated with the presence of thermotolerant coliforms, though not with fecal streptococci.⁹ It is also possible that in settings with high population density, the mere presence/absence of a latrine is an inadequate metric, and latrine density within a given radius may be a more nuanced measure. Latrine density as well as population density within up to 100 m of a tubewell have been shown to be correlated with microbiological tubewell water quality in rural Bangladesh.²³ In contrast, a study in Kenya found no association between pit latrine density within 100 m and thermotolerant coliforms in hand-dug wells; however, latrine density was associated with other contami-

nants suggesting infiltration from latrines, such as nitrate and chloride.¹²

Older wells were more likely to have drainage channels and concrete platforms, but these were more likely to be deteriorated (e.g., broken or eroded), and older wells had higher sanitary scores overall. These could reflect secular changes in well construction practices, deterioration of tubewell components over time, or a combination of both. Surprisingly, older wells were also less likely to be located within 10 m of a latrine. This could be due to increasing population density forcing latrines and wells to be built closer to each other due to space constraints.

Season modified the associations between sanitary risk factors and tubewell water quality. This suggests that different contamination mechanisms dominate during wet conditions with elevated groundwater levels and periodic washing away of fecal contamination from child and animal feces by rain versus dry conditions with relatively lower groundwater tables and conditions that enable accumulation of fecal material near the wellhead. Previous research also suggests seasonal variations in risk factors for tubewell water quality.²³

Taken together with previous evidence from other settings, our findings suggest that localized short-circuiting of surface contamination from the wellhead catchment into the well is a more important route for contamination of shallow groundwater sources than subsurface transport of pathogens.^{9–11,21} Efforts to improve the microbiological quality of shallow groundwater sources should therefore include adequate construction and maintenance of wellhead components. Additionally, groundwater sources are known to become further contaminated during storage in the home.^{6,18,30} Efforts to reduce waterborne fecal exposures should therefore include measures to minimize point-of-use contamination. When tracking progress toward global safe water access targets, water quality measured in the household should be considered in addition to source type, relevant sanitary risk factors and water quality measured at the source.

Limitations. Our study setting represents unique conditions that are not representative of shallow groundwater sources globally. Bangladesh is the most densely populated country in the world and is characterized by a high groundwater table.²⁵ These conditions, combined with the common use of pit latrines for on-site sanitation with typically unlined pits, would theoretically present enabling conditions for infiltration of pathogens from latrines into groundwater aquifers, such that the lack of a harmful association between latrine presence and water quality in this worst-case setting would be expected to hold in lower-risk environments. However, Bangladesh is dominated by alluvial sediments and the area encompassing our study site, Mymensingh, is categorized as recent flood plains consisting of grey clay, silt, and fine sand.²⁵ Microorganism fate and transport will differ in settings with different lithology. It is possible that the soil characteristics in our study area provided an effective barrier against pathogen transport in the subsurface; a study in India detected fecal coliforms and nitrate in wells near latrines in settings with fractured rock but not with alluvial formations.³¹ This is consistent with the observation that tubewells in rural Bangladesh have relatively good water quality despite the high population density and reliance on pit latrines; only 14% of the wells we tested exceeded the WHO low-risk threshold of 10 CFU/100 mL.²⁷ It is possible

that this limited our ability to detect associations, and settings with different soil characteristics and heavier contamination could reveal a stronger relationship between water quality and sanitary inspection scores. However, our findings are broadly consistent with previous evidence from a variety of settings.

Other limitations pertain to inherent deficiencies of water quality testing. Microbiological water quality is highly temporally variable, with significant fluctuations even on an hourly basis.³² While our sampling scheme addressed seasonal variability between the rainy and dry seasons, ultimately, grab samples cannot adequately capture transient spikes of contamination, such as those during or immediately after rainfall. Additionally, while *E. coli* is widely used as a fecal indicator organism and has been shown to be associated with waterborne illness risk,³³ it poorly correlates with the presence of actual pathogens in water.³⁴ In the subsurface, *E. coli* has significantly shorter survival than viruses, and depending on the soil pore size, it can be more effectively filtered by soil particles than viruses and less effectively filtered than protozoa.³⁵ *Escherichia coli* can also be naturally present and grow in tropical soils.^{36,37} These sources of potential misclassification of water quality would weaken any observed relationship between sanitary inspections and water quality as measured by *E. coli* in infrequent grab samples. Alternative water quality indicators such as conductance, nitrate, or chloride may aid in detecting fecal input into water sources.

Finally, a limitation of the sanitary survey tool is its failure to account for local groundwater and geological conditions. Characteristics of local aquifers, the depth of the groundwater profile, and soil type and strata significantly impact the subsurface transport of contaminants. It has been suggested that blanket wellhead protection areas based on conceptual groundwater models fail to account for the variability in pathogen fate and transport in the subsurface,³⁵ and a range of latrine siting guidelines with respect to proximity to drinking water sources have been recommended in different settings.³⁸ Augmenting sanitary inspections with an assessment of the groundwater table and characterization of local soil type could help elucidate the relationships between sanitary risk scores and water quality.

Conclusions. Identifying the risk factors leading to groundwater contamination would allow targeted corrective action to reduce waterborne fecal exposure. It would also improve the tracking of progress toward global targets for safe water provision by augmenting the commonly used source type-based definition of improved versus unimproved sources with identification of relevant sanitary risk factors to provide a more accurate and nuanced assessment of the safety of a given water source. While the use of sanitary inspection scores has been recommended to assess the safety of water sources, our analysis suggests that the sanitary score was unassociated with *E. coli* presence and concentration in shallow tubewells in Bangladesh. Individual risk factors measuring the integrity of the wellhead were associated with water quality but had no predictive power for the presence of *E. coli*, whereas the presence of a latrine within 10 m of a well was not associated with water quality in the wet season and was associated with less frequent *E. coli* detection during the dry season. These findings indicate that observed characteris-

tics of a tubewell, as measured by sanitary inspections in their current form, do not sufficiently characterize the microbiological quality of tubewell water, as measured by *E. coli*. Assessment of local groundwater and geological conditions as part of the sanitary inspection and improved water quality indicators may reveal more clear relationships. Our findings also suggest that the dominant contamination route for shallow groundwater is short-circuiting at the wellhead, and efforts to improve the safety of groundwater sources should include adequate wellhead construction and maintenance.

Received June 16, 2016. Accepted for publication November 27, 2016.

Published online January 23, 2017.

Note: Supplemental tables appear at www.ajtmh.org.

Acknowledgments: icddr,b acknowledges with gratitude the commitment of USAID to its research efforts. icddr,b is also grateful to the Governments of Bangladesh, Canada, Sweden, and the United Kingdom for providing core/unrestricted support.

Financial support: This work was supported by the U.S. Agency for International Development (USAID).

Disclaimer: The contents of this manuscript are the sole responsibility of the icddr,b and do not necessarily reflect the views of USAID, CDC, or the U.S. Government. USAID had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' addresses: Ayse Ercumen, Benjamin F. Arnold, and John M. Colford Jr., Division of Epidemiology, School of Public Health, University of California, Berkeley, CA, E-mails: aercumen@berkeley.edu, benarnold@berkeley.edu, and jcolford@berkeley.edu. Abu Mohd Naser, Rollins School of Public Health, Emory University, Atlanta, GA, E-mail: abu.mohd.naser.titu@emory.edu. Leanne Unicomb, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh, E-mail: leanne@icddr.org. Stephen P. Luby, School of Medicine, Stanford University, Stanford, CA, E-mail: sluby@stanford.edu.

REFERENCES

1. World Health Organization/United Nations Children's Fund Joint Water Supply and Sanitation Monitoring Programme, 2015. *Progress on Sanitation and Drinking Water*. Geneva, Switzerland: WHO.
2. Bain R, Cronk R, Wright J, Yang H, Slaymaker T, Bartram J, 2014. Fecal contamination of drinking-water in low- and middle-income countries: a systematic review and meta-analysis. *PLoS Med* 11: e1001644.
3. Onda K, LoBuglio J, Bartram J, 2012. Global access to safe water: accounting for water quality and the resulting impact on MDG progress. *Int J Environ Res Public Health* 9: 880–894.
4. United Nations Economic and Social Council, 2016. *Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators*. United Nations Economic and Social Council. Available at: <http://unstats.un.org/unsd/statcom/47th-session/documents/2016-2-IAEG-SDGs-E.pdf>. Accessed June 6, 2016.
5. Bain R, Cronk R, Hossain R, Bonjour S, Onda K, Wright J, Yang H, Slaymaker T, Hunter P, Prüss-Ustün A, Bartram J, 2014. Global assessment of exposure to faecal contamination through drinking water based on a systematic review. *Trop Med Int Health* 19: 917–927.
6. Shields KF, Bain RE, Cronk R, Wright JA, Bartram J, 2015. Association of supply type with fecal contamination of source water and household stored drinking water in developing countries: a bivariate meta-analysis. *Environ Health Perspect* 123: 1222–1231.
7. Luby SP, Gupta SK, Sheikh MA, Johnston RB, Ram PK, Islam MS, 2008. Tubewell water quality and predictors of

- contamination in three flood-prone areas in Bangladesh. *J Appl Microbiol* 105: 1002–1008.
8. Lloyd BJ, Bartram JK, 1991. Surveillance solutions to microbiological problems in water quality control in developing countries. *Water Sci Technol* 24: 61–75.
 9. Howard G, Pedley S, Barrett M, Nalubega M, Johal K, 2003. Risk factors contributing to microbiological contamination of shallow groundwater in Kampala, Uganda. *Water Res* 37: 3421–3429.
 10. Cronin AA, Breslin N, Gibson J, Pedley S, 2006. Monitoring source and domestic water quality in parallel with sanitary risk identification in northern Mozambique to prioritise protection interventions. *J Water Health* 4: 333–345.
 11. Godfrey S, Timo F, Smith M, 2006. Microbiological risk assessment and management of shallow groundwater sources in Lichinga, Mozambique. *Water Environ J* 20: 194–202.
 12. Wright JA, Cronin A, Okotto-Okotto J, Yang H, Pedley S, Gundry SW, 2012. A spatial analysis of pit latrine density and groundwater source contamination. *Environ Monit Assess* 185: 4261–4272.
 13. Islam MS, Siddika A, Khan MN, Goldar MM, Sadique MA, Kabir A, Huq A, Colwell RR, 2001. Microbiological analysis of tube-well water in a rural area of Bangladesh. *Appl Environ Microbiol* 67: 3328–3330.
 14. Luby S, Islam MS, Johnston R, 2006. Chlorine spot treatment of flooded tube wells, an efficacy trial. *J Appl Microbiol* 100: 1154–1158.
 15. Leber J, Rahman MM, Ahmed KM, Mailloux B, Geen A, 2011. Contrasting influence of geology on *E. coli* and arsenic in aquifers of Bangladesh. *Ground Water* 49: 111–123.
 16. van Geen A, Ahmed KM, Akita Y, Alam MJ, Culligan PJ, Emch M, Escamilla V, Feighery J, Ferguson AS, Knappett P, Layton AC, Mailloux BJ, McKay LD, Mey JL, Serre ML, Streatfield PK, Wu J, Yunus M, 2011. Fecal contamination of shallow tubewells in Bangladesh inversely related to arsenic. *Environ Sci Technol* 45: 1199–1205.
 17. Ferguson AS, Layton AC, Mailloux BJ, Culligan PJ, Williams DE, Smartt AE, Saylor GS, Feighery J, McKay LD, Knappett PSK, Alexandrova E, Arbit T, Emch M, Escamilla V, Ahmed KM, Alam MJ, Streatfield PK, Yunus M, van Geen A, 2012. Comparison of fecal indicators with pathogenic bacteria and rotavirus in groundwater. *Sci Total Environ* 431: 314–322.
 18. Ercumen A, Naser AM, Unicomb L, Arnold BF, Jr JMC, Luby SP, 2015. Effects of source-versus household contamination of tubewell water on child diarrhea in rural Bangladesh: a randomized controlled trial. *PLoS One* 10: e0121907.
 19. Knappett PSK, McKay LD, Layton A, Williams DE, Alam MJ, Huq MR, Mey J, Feighery JE, Culligan PJ, Mailloux BJ, Zhuang J, Escamilla V, Emch M, Perfect E, Saylor GS, Ahmed KM, van Geen A, 2011. Implications of fecal bacteria input from latrine-polluted ponds for wells in sandy aquifers. *Environ Sci Technol* 46: 1361–1370.
 20. Knappett PSK, Escamilla V, Layton A, McKay LD, Emch M, Williams DE, Huq R, Alam J, Farhana L, Mailloux BJ, Ferguson A, Saylor GS, Ahmed KM, van Geen A, 2011. Impact of population and latrines on fecal contamination of ponds in rural Bangladesh. *Sci Total Environ* 409: 3174–3182.
 21. Knappett PSK, McKay LD, Layton A, Williams DE, Alam MJ, Mailloux BJ, Ferguson AS, Culligan PJ, Serre ML, Emch M, Ahmed KM, Saylor GS, van Geen A, 2012. Unsealed tubewells lead to increased fecal contamination of drinking water. *J Water Health* 10: 565.
 22. Ferguson AS, Mailloux BJ, Ahmed KM, Geen AV, McKay LD, Culligan PJ, 2011. Hand-pumps as reservoirs for microbial contamination of well water. *J Water Health* 9: 708–717.
 23. Escamilla V, Knappett PSK, Yunus M, Streatfield PK, Emch M, 2013. Influence of latrine proximity and type on tubewell water quality and diarrheal disease in Bangladesh. *Ann Assoc Am Geogr* 103: 299–308.
 24. U.S. Environmental Protection Agency (USEPA), 2002. *Method 1604: Total Coliforms and Escherichia coli in Water by Membrane Filtration Using a Simultaneous Detection Technique (MI Medium)*. Washington, DC: U.S. Environmental Protection Agency.
 25. Ahmed KM, Bhattacharya P, Hasan MA, Akhter SH, Alam SMM, Bhuyian MAH, Imam MB, Khan AA, Sracek O, 2004. Arsenic enrichment in groundwater of the alluvial aquifers in Bangladesh: an overview. *Appl Geochem* 19: 181–200.
 26. McNutt LA, Wu C, Xue X, Hafner JP, 2003. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am J Epidemiol* 157: 940–943.
 27. World Health Organization, 1997. *Guidelines for Drinking-Water Quality*, 2nd edition. Geneva, Switzerland: WHO.
 28. Santamaria J, Toranzos GA, 2003. Enteric pathogens and soil: a short review. *Int Microbiol* 6: 5–9.
 29. Verheyen J, Timmen-Wego M, Laudien R, Boussaad I, Sen S, Koc A, Uesbeck A, Mazou F, Pfister H, 2009. Detection of adenoviruses and rotaviruses in drinking water sources used in rural areas of Benin, west Africa. *Appl Environ Microbiol* 75: 2798–2801.
 30. Wright J, Gundry S, Conroy R, 2004. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. *Trop Med Int Health* 9: 106–117.
 31. Pujari PR, Padmakar C, Labhasetwar PK, Mahore P, Ganguly AK, 2012. Assessment of the impact of on-site sanitation systems on groundwater pollution in two diverse geological settings: a case study from India. *Environ Monit Assess* 184: 251–263.
 32. Levy K, Hubbard AE, Nelson KL, Eisenberg JNS, 2009. Drivers of water quality variability in northern coastal Ecuador. *Environ Sci Technol* 43: 1788–1797.
 33. Gruber JS, Ercumen A, Jr JMC, 2014. Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PLoS One* 9: e107429.
 34. Wu J, Long SC, Das D, Dörner SM, 2011. Are microbial indicators and pathogens correlated? A statistical analysis of 40 years of research. *J Water Health* 9: 265–278.
 35. Taylor R, Cronin A, Pedley S, Barker J, Atkinson T, 2004. The implications of groundwater velocity variations on microbial transport and wellhead protection: review of field evidence. *FEMS Microbiol Ecol* 49: 17–26.
 36. Hardina CM, Fujioka RS, 1991. Soil: the environmental source of *Escherichia coli* and enterococci in Hawaii's streams. *Environ Toxicol Water Qual* 6: 185–195.
 37. Fujioka R, Sian-Denton C, Borja M, Castro J, Morphew K, 1998. Soil: the environmental source of *Escherichia coli* and enterococci in Guam's streams. *J Appl Microbiol* 85: 83S–89S.
 38. Graham JP, Polizzotto ML, 2013. Pit latrines and their impacts on groundwater quality: a systematic review. *Environ Health Perspect* 121: 521–530. Available at: http://hsr.himmelfarb.gwu.edu/sphhs_enviro_facpubs/36/. Accessed July 21, 2016.