**Original Paper** 

# Association Among Household Water, Sanitation, and Hygiene (WASH) Status and Typhoid Risk in Urban Slums: Prospective Cohort Study in Bangladesh

Birkneh Tilahun Tadesse<sup>1\*</sup>, MD, PhD; Farhana Khanam<sup>2\*</sup>, MBBS; Faisal Ahmmed<sup>2</sup>, MSc; Xinxue Liu<sup>3</sup>, PhD; Md Taufiqul Islam<sup>2</sup>, MBBS; Deok Ryun Kim<sup>1</sup>, PhD; Sophie SY Kang<sup>1</sup>, MSc; Justin Im<sup>1</sup>, PhD; Fahima Chowdhury<sup>2</sup>, MBBS; Tasnuva Ahmed<sup>2</sup>, MBBS; Asma Binte Aziz<sup>1</sup>, MBBS; Masuma Hoque<sup>2</sup>, MBBS; Juyeon Park<sup>1</sup>, MSc; Gideok Pak<sup>1</sup>, MSc; Hyon Jin Jeon<sup>1</sup>, MSc; Khalequ Zaman<sup>2</sup>, PhD; Ashraful Islam Khan<sup>2</sup>, PhD; Jerome H Kim<sup>1</sup>, MD; Florian Marks<sup>1,4,5</sup>, PhD; Firdausi Qadri<sup>2</sup>, PhD; John D Clemens<sup>1,2</sup>, MD

<sup>5</sup>Madagascar Institute for Vaccine Research, University of Antananarivo, Antananarivo, Madagascar

\*these authors contributed equally

#### **Corresponding Author:**

Birkneh Tilahun Tadesse, MD, PhD Epidemiology, Public Health, Impact Unit International Vaccine Institute SNU Research Park 1 Gwanak-ro, Gwanak-gu Seoul, 08826 Republic of Korea Phone: 82 1098041348 Email: birkneh.tadesse@ivi.int

# Abstract

**Background:** Typhoid fever, or enteric fever, is a highly fatal infectious disease that affects over 9 million people worldwide each year, resulting in more than 110,000 deaths. Reduction in the burden of typhoid in low-income countries is crucial for public health and requires the implementation of feasible water, sanitation, and hygiene (WASH) interventions, especially in densely populated urban slums.

**Objective:** In this study, conducted in Mirpur, Bangladesh, we aimed to assess the association between household WASH status and typhoid risk in a training subpopulation of a large prospective cohort (n=98,087), and to evaluate the performance of a machine learning algorithm in creating a composite WASH variable. Further, we investigated the protection associated with living in households with improved WASH facilities and in clusters with increasing prevalence of such facilities during a 2-year follow-up period.

**Methods:** We used a machine learning algorithm to create a dichotomous composite variable ("Better" and "Not Better") based on 3 WASH variables: private toilet facility, safe drinking water source, and presence of water filter. The algorithm was trained using data from the training subpopulation and then validated in a distinct subpopulation (n=65,286) to assess its sensitivity and specificity. Cox regression models were used to evaluate the protective effect of living in "Better" WASH households and in clusters with increasing levels of "Better" WASH prevalence.

**Results:** We found that residence in households with improved WASH facilities was associated with a 38% reduction in typhoid risk (adjusted hazard ratio=0.62, 95% CI 0.49-0.78; *P*<.001). This reduction was particularly pronounced in individuals younger than 10 years at the first census participation, with an adjusted hazard ratio of 0.49 (95% CI 0.36-0.66; *P*<.001). Furthermore, we observed an inverse relationship between the prevalence of "Better" WASH facilities in clusters and the incidence of typhoid,

RenderX

<sup>&</sup>lt;sup>1</sup>Epidemiology, Public Health, Impact Unit, International Vaccine Institute, Seoul, Republic of Korea

<sup>&</sup>lt;sup>2</sup>Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh

<sup>&</sup>lt;sup>3</sup>Oxford Vaccine Group, Department of Pediatrics, University of Oxford, Oxford, United Kingdom

<sup>&</sup>lt;sup>4</sup>Cambridge Institute of Therapeutic Immunology and Infectious Disease, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom

although this association was not statistically significant in the multivariable model. Specifically, the adjusted hazard of typhoid decreased by 0.996 (95% CI 0.986-1.006) for each percent increase in the prevalence of "Better" WASH in the cluster (P=.39).

**Conclusions:** Our findings demonstrate that existing variations in household WASH are associated with differences in the risk of typhoid in densely populated urban slums. This suggests that attainable improvements in WASH facilities can contribute to enhanced typhoid control, especially in settings where major infrastructural improvements are challenging. These findings underscore the importance of implementing and promoting comprehensive WASH interventions in low-income countries as a means to reduce the burden of typhoid and improve public health outcomes in vulnerable populations.

(JMIR Public Health Surveill 2023;9:e41207) doi: 10.2196/41207

#### **KEYWORDS**

water; sanitation; sanitary; contaminated; contamination; hygiene; hygienic; WASH; water, sanitation and hygiene; typhoid fever; enteric fever; typhus; typhoid; enteric; salmonella; protection; recursive partitioning; Bangladesh; low- and middle-income countries; LMIC; bacteria; bacterial; bacterial infection; machine learning; algorithm; algorithms; model; low income; slum; slums; risk; infection control; incidence; prevalence; epidemiology; epidemiological; poverty

# Introduction

Typhoid fever, also known as enteric fever, is a deadly febrile illness caused by the Salmonella enterica serovar Typhi (Salmonella Typhi) [1,2]. Based on 2019 data, approximately 9 million people contract typhoid annually, leading to 110,000 deaths [3]. Common symptoms include persistent fever, fatigue, headache, nausea, abdominal pain, and digestive issues, and occasionally a rash [2,3]. Severe cases of typhoid fever can have life-threatening complications and result in fatality [4]. The effectiveness of treating typhoid has been compromised by the growing resistance to various antibiotics [5,6]. To address this challenge, the World Health Organization advocates for the use of the typhoid conjugate vaccine (TCV) in children aged 6 months and older [3]. Nevertheless, the implementation of vaccination programs alone will not be enough to effectively control and eliminate the transmission of typhoid fever, particularly in urban slum areas where water, sanitation, and hygiene (WASH) facilities and practices are inadequate [7].

The primary mode of transmission for *Salmonella* Typhi is through the fecal-oral route via consumption of food and water that has been contaminated [2]. Consequently, inadequate access to clean water, improved sanitary facilities, appropriate waste management, and proper hand hygiene practices can contribute to an increased risk of typhoid fever [7,8]. For instance, successful elimination of typhoid fever in high-income countries is primarily attributed to major improvements in WASH facilities and practices [7,9]. However, substantial WASH modernization requires major financial, capital, and political commitment, and continues to be an unmet need in resource-constrained settings where unsafe drinking water and lack of sanitary facilities, augmented by unplanned urbanization, significantly contribute to increased typhoid transmission [10,11].

The need for multidimensional interventions, which now include vaccination using available TCVs, to prevent exposure to and transmission of typhoid fever demands implementation of feasible approaches to improving WASH, especially within densely populated urban settings. The emphasis on implementing practical enhancements in WASH is significant, yet remains relevant until a comprehensive upgrade of WASH practices and infrastructure is implemented in settings with high

```
https://publichealth.jmir.org/2023/1/e41207
```

typhoid burden. Addressing the problem of typhoid fever in urban slums is particularly important as the highest burden has been reported in these settings including the occurrence of enhanced environmental transmission [12,13]. For example, in urban slums in Bangladesh, a high burden of typhoid fever (incidence of 200 per 100,000 person years of observation; PYO) and paratyphoid fever (incidence of 40 per 100,000 PYO) was reported [14]. The implementation of WASH practices has the potential to yield critical benefits, including the reduction of *Salmonella* Typhi transmission within communities. By mitigating the prevalence of this bacterium, the burden of bloodstream infections caused by *Salmonella* Typhi can be diminished. As a result, the consumption of antibiotics may be significantly reduced, contributing to the ongoing battle against antimicrobial resistance.

Our earlier work demonstrated that existing improvements in WASH facilities in a Kolkata slum were associated with reduced risk of typhoid (hazard ratio [HR]=0.57, 95% CI (0.37-0.90; P=.015) [8]. In this work, we developed a decision tree defining "Good" versus "Not Good" household WASH using 4 variables: source of drinking water, hand washing with soap, presence of flush toilet, and source of daily water use. The decision tree successfully discriminated between households at higher versus lower risk of typhoid (area under the curve [AUC] of 58%, 95% CI 54-61; sensitivity of 90.4%, 95% CI 84.2-94.8; and specificity of 19.8%, 95.5 CI 19.3-20.3). Notably, though, the data used for this analysis were accrued more than 20 years ago, possibly limiting their generalizability to circumstances in impoverished urban slums with endemic typhoid today.

Here we investigate whether existing improvements in WASH in a contemporary slum in Dhaka were also associated with a reduced risk of typhoid. We reexamined the hypothesis that variations in household WASH already present in urban slums can successfully predict the risk of typhoid fever. We test this hypothesis using the data from prospective, population-based, surveillance for typhoid fever in the control clusters of a cluster randomized trial of TCV recently completed in urban slums of Mirpur, Dhaka, Bangladesh [15].

XSL•FO RenderX

# Methods

# Cluster Randomized Trial on the Vi-Tetanus Toxoid Vaccine

A cluster randomized trial (CRT) was conducted between May 2018 and March 2020 with the aim to evaluate the protection, including herd protection, conferred by Vi-tetanus toxoid conjugate vaccine (Vi-TT, Typbar-TCV, Bharat Biotech International Limited) when administered in a mass vaccination campaign. This participant- and observer-blinded CRT included children 9 months to <16 years within a censused total population of 205,760, where 150 clusters were randomized to receive Vi-TT or Japanese encephalitis (JE) vaccines (control clusters) in a 1:1 ratio. These clusters were approximately of comparable geographic size with mean area of 0.02 (SD 0.01) km<sup>2</sup> and included contiguous households with approximately equal total populations separated by natural borders wherever possible [7]. Age-eligible, nonpregnant residents who provided informed consent received a single dose of the assigned vaccine. The CRT was implemented in the Mirpur area in Bangladesh located in the north-east of Dhaka city. Mirpur is a densely populated settlement with a total surface area of 58.66 km<sup>2</sup> and approximately 3.5 million residents, serving as a field study site since 1987.

A baseline census was performed between February 14 and March 25, 2018, to enumerate the entire population in this study's area, except nonpermanent residents and individuals who planned to move out within a month. During the census, individual- and household-level socioeconomic and demographic data, details on water-sanitation-hygiene facilities and practices, and geo-positioning coordinates of each household were collected after obtaining informed consent. A "household" was defined as persons sharing the same cooking pot. The census was subsequently updated at approximately 6-month intervals to capture all births, deaths, and migrations. At each census, 13 nonbinary variables characterizing household WASH were collected. Because there were no significant differences between the household WASH status at first census and the subsequent updates by household, WASH characteristics at first census participation were used for this analysis. Accordingly, WASH variables collected during the baseline census were used to characterize household WASH status for individuals present in this study's area during the baseline census; the date of participation in the baseline census was taken as the start date of follow-up. For individuals who moved into the clusters after the baseline census and for births, the WASH variables collected during the first census update after the onset of their first participation were used to characterize household WASH status. The date of migration in and the date of birth were considered as the starting date of follow-up of these participants.

#### **Typhoid Fever Surveillance**

Surveillance for typhoid fever started on February 26, 2018, in all age groups using a common protocol across all 8 clinical facilities (Mirpur Field Clinic; Suhrawardy Hospital; Radda Barnen; Adhunik Hospital; Shishu hospital; Kurmitola General Hospital; Kingston Hospital; Islami Bank Hospital) serving this study's population. Biweekly home visits of the entire

```
https://publichealth.jmir.org/2023/1/e41207
```

population were performed by community health workers to promote care seeking of all household members for febrile illnesses at one of this study's facilities. Blood samples for microbiological culture (3 mL for  $\leq$ 17 years, and 5 mL for >17 years old) were collected for individuals presenting with history of fever  $\geq 2$  days or with axillary temperature of  $\geq 38$  °C, and clinical findings were systematically recorded by a study physician. Identity cards distributed during the censuses enabled identification of participants when presenting for febrile illness related care and treatment. For individuals not having the identification cards, computerized censuses on electronic tablets available at each study site were used to confirm identity. Blood specimens for culture were obtained and transported on the same day of collection to laboratories at the icddr,b. Blood cultures were monitored in an automated BacT/ALERT system, and subculturing was done on MacConkey, chocolate, and blood agar plates following a positive signal by the machine. Colony morphology, microscopic examination of Gram stain, biochemical tests, and slide agglutination tests were used for identification of Salmonella isolates. Salmonella-specific antisera including polyvalent O and O1, serogroup "O" Salmonella antigens [D], and flagellar "H" antigens [d] (Denka Sieken) were used for agglutination [16,17]. Treatment for patients with typhoid fever was based on the physicians' discretion, guided by antimicrobial susceptibility results. A single febrile episode was considered for febrile visits with the onset of fever ≤14 days after discharge for a previous febrile visit. A home visit was performed within 14 days of discharge for all participants with positive blood cultures to confirm that the person whose name was given on the day of presentation to the treatment center had indeed sought treatment on that date. A typhoid fever episode was defined as a febrile episode with a confirmed identity during which at least 1 blood culture yielded S enterica serotype Typhi (Salmonella Typhi).

### Variable Selection to Define "Better" Versus "Not Better" WASH

As a first step, we partitioned each WASH variable collected during the censuses into a binary variable defining "Better" versus "Not Better" household WASH. The variables were partitioned based on substantive judgment aligned with the local context, and without prior knowledge of the typhoid incidence rates in participants with different categories of these variables. We then divided the households of total population in the JE clusters (75 clusters) at random into 2 mutually exclusive subpopulations: 98,087 (60%) of households constituting a "training" subpopulation and the other 65,286 (40%) a "validation" subpopulation. Variables associated with typhoid risk in bivariate Cox regression models at  $P \le .20$  in the training subpopulation were selected for model development. Inclusion of all 13 WASH variables regardless of associations of the individual variables with the risk of typhoid, did not reveal any improvement in model performance. Proportionality assumptions for each variable were confirmed before inclusion in the bivariate models.

```
XSL•FO
RenderX
```

# **Development of the WASH Decision Tree to Predict Typhoid in the Training Subpopulation**

As described elsewhere, we developed a composite WASH variable based on the selected candidate WASH variables to predict typhoid risk using a recursive partitioning-based algorithm in the training subpopulation [8,18]. In brief, we developed a binary WASH rule that predicts typhoid fever risk using individual WASH variables that were individually associated with the risk of typhoid fever and then conjoined as specified by the decision tree. Ultimately, the decision tree represents a composite WASH rule based on a combination of these WASH variables (Better versus Not Better WASH) predicting a classification of no typhoid versus typhoid, respectively. The decision tree then yielded several terminal nodes predictive of "higher" or "lower" risk of typhoid fever, which then served as alternative cutoffs used to create a dichotomous composite rule for "higher" or "lower" risk of typhoid, using the Youden index [12]. The WASH features in the decision tree predictive of "higher" risk of typhoid in this dichotomous typhoid variable then defined "Not Better" WASH and the features predictive of "lower" risk of typhoid in the dichotomous variable defined "Better" WASH.

The prediction rule was designed to discriminate the risk of typhoid fever in a dynamic population over the 2-year period of postvaccination follow-up in the trial. The algorithm was first evaluated in the training subpopulation assuming 1:1 ratio for a default loss of function of the cost of false positive and false negative classifications, with at least 300 observations at each terminal node. Subsequently, the algorithm was cross validated with estimation of the cross-validation error in 1 of 10 randomly assembled partitions of the training subpopulation. The 10-fold cross-validation technique was employed with the "rpart" package in the R software (R Foundation for Statistical Computing). We first constructed a tree using the training data, including terminal nodes. Then, we pruned the tree to obtain the smallest tree with the lowest misclassification error. The minimal complexity parameter was used to prune the model in finding the optimal prediction rule providing the minimum error with at least 2 terminal nodes in the tree. A receiver operating characteristic (ROC) curve was used to define the AUC for different cutoffs for dichotomizing the predictive rule, with the optimal cutoff for typhoid prediction selected as that which maximized the Youden Index, which includes sensitivity plus specificity, in relation to the proportion of participants in the terminal nodes developing typhoid fever [19].

# Evaluation of the Composite WASH Variable in a Validation Subpopulation and Prediction of Typhoid Risk in the Total Population

To evaluate reproducibility of sensitivity and specificity of the composite WASH variable developed in the training subpopulation, we tested the variable in a distinct *validation* subpopulation. After ascertaining concordance of sensitivity (the proportion of participants developing typhoid fever who lived in households with "Not Better" WASH) and specificity (the proportion of participants not developing typhoid fever who lived in households with "Better" WASH) in predicting typhoid risk between the training and validation subpopulations,

```
https://publichealth.jmir.org/2023/1/e41207
```

XSI•F(

we then evaluated the ability of the variable to predict typhoid risk in the total population in the JE clusters.

## **Data Analysis**

We constructed Cox proportional hazard regression models to evaluate the association between household WASH status and risk of typhoid in household inhabitants as a function of time to the first typhoid episode. Follow-up of persons in the baseline census began at the time of the census, and follow-up of those entering later as births or in-migrants began on the date of entry into the population. Follow-up was right censored by death, outmigration, and the end of this study. Before inclusion in the models, proportionality assumptions for all independent variables were assessed by a bivariate Cox model. The models were adjusted for randomization stratifying variables including geographic ward, longer than median distance to this study's clinics, and number of eligible children residing in each cluster at baseline. As well, other baseline covariates, including age in years, number of participants in clusters, religion (Muslim), and monthly household expenditure (in Bangladeshi taka), were fitted as independent variables if they were associated with the hazard of typhoid in bivariate Cox models at P<.15. The average expenditure of households in a cluster was used as the monthly expenditure of households with unspecified monthly expenditure. The model coefficient for living in a "Better" WASH household was exponentiated to estimate the HR of typhoid fever, and the 95% CI for the HR was estimated using a robust sandwich method to adjust for the design effect of the clusters. Protection by Better WASH against typhoid fever was assessed as  $[(1 - HR) \times 100\%)$  with corresponding 95% interval limits for percentage protection. We also used Cox models to evaluate the protective association between the cluster prevalence of "Better" WASH households and the risk of typhoid in the total population during the follow-up period. The "Better" WASH cluster prevalence during the 2 years follow-up was calculated as the proportion of PYO contributed by members of households with "Better" WASH in the cluster divided by the total PYO for members of all households in the same cluster. We also assessed the protective association between increasing levels of "Better" WASH cluster prevalence, expressed in Cox models dimensionally, and the risk of typhoid after adjustment for the above-mentioned covariates. In these models, "Better" WASH coverage of the cluster was assigned as the "Better" WASH coverage datapoint for each resident in the respective cluster.

The analysis was performed using the *rpart* package for decision tree modelling, *rpart.plot* package for tree plotting, *pROC* package for the ROC curve, *survival* package for the Cox model, and *dplyr* package for data management under R Studio analytical software (R Foundation for Statistical Computing) [20-22]. The presence of collinearity in the multivariate models was assessed by inspection of variance inflation factors, with high variance inflation factors indicative of collinearity. For all the statistical analyses, a P < .05 (2-tailed) was taken as the margin of statistical significance.

#### **Ethical Considerations**

The trial has received ethical approval from the Research and Ethical Review Committees of icddr,b (PR-17115), as well as

the institutional review boards of Oxford University and the University of Maryland. Informed written consent from all participants was obtained, with parental or guardian consent for those younger than the age of 16 years, and assent from participants aged 11 years to younger than 16 years. The trial was monitored by a local Data and Safety Monitoring Board established by icddr,b, as well as an international Data and Safety Monitoring Board established by the University of Maryland. This study's data are anonymous and were securely stored in access-controlled cabinets. The trial has been registered in the ISRCTN Registry under the identifier ISRCTN11643110.

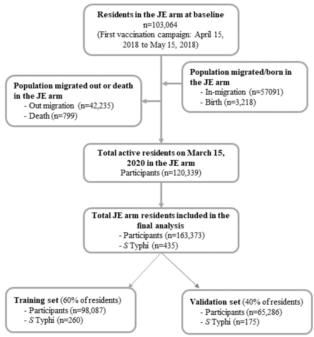
# Results

# Assembly and Characteristics of Training and Validation Subpopulations

A total of 103,064 individuals living in 25,478 households within the control clusters for the trial were enumerated during

the baseline census. In these households, 51,211 (49.7%) individuals were men, and the mean age was 26.8 (SD 17.3) years. During the 2 years of follow-up, 57,091 persons migrated into this study's area, 3218 were born, and 43,034 migrated out of this study's area or died, making the population residing in the JE clusters at the last census 120,339 persons, and the population ever followed in these clusters during the 2 years 163,373 persons (Figure 1). The training subpopulation comprised 98,087 individuals with a mean age 25.9 (SD 17.3) years, while the validation subpopulation included 65,286 individuals with a mean age 25.9 (SD 17.3) years (Table 1). Men were 48,638 (49.6%) in the training subpopulation and 32,496 (49.8%) in the validation subpopulation. During the 2-years of follow-up, a total of 435 first episodes of typhoid cases were detected, yielding an incidence of 209/100,000 PYO in the total population residing in the control clusters. The numbers of typhoid cases in the training and validation subpopulations were 260 and 175, respectively.

**Figure 1.** CONSORT (Consolidated Standards of Reporting Trials) diagram of population dynamics during April 15, 2018, and March 15, 2020. JE: Japanese encephalitis; N: number; *S* Typhi: *Salmonella enterica* serovar Typhi.





**Table 1.** Baseline characteristics of the total, training, and validation populations in the control clusters of the Vi-TT<sup>a</sup> cluster randomized trial, Dhaka, Bangladesh.

Characteristics	Full set	Training set	Validation set
Age (years), mean (SD)	25.9 (17.3)	25.9 (17.3)	25.9 (17.3)
Gender: men, n (%)	81,134 (49.7)	48,638 (49.6)	32,496 (49.8)
Religion: Muslim, n (%)	161,591 (98.9)	97,004 (98.9)	64,587 (98.9)
Monthly expenditure (BDT <sup>b</sup> ) <sup>c</sup> , mean (SD)	17,420.3 (9053.2)	17,397.5 (7793)	17,454.6 (10,670.3)
Toilet facility: private, n (%)	68838 (42.1)	41155 (42)	27683 (42.4)
Adult toilet: flush toilet, n (%)	7804 (4.8)	4617 (4.7)	3187 (4.9)
Child toilet: flush toilet, n (%)	1000 (0.6)	580 (0.6)	420 (0.6)
Source of drinking water: private tap, well, or pump; water vendor, n (%)	63,263 (38.7)	37,854 (38.6)	25,409 (38.9)
Freated drinking water, n (%)	122,258 (74.8)	73,527 (75)	48,731 (74.6)
Freated cleaning water, n (%)	475 (0.3)	300 (0.3)	175 (0.3)
WASH <sup>d</sup> before meals, n (%)	121,595 (74.4)	73,019 (74.4)	48,576 (74.4)
Hand wash after defecation, n (%)	159,390 (97.6)	95,698 (97.6)	63,692 (97.6)
Handwashing water available in HH <sup>e</sup> , n (%)	159,193 (97.5)	95,616 (97.5)	63,577 (97.4)
Handwashing soap available in HH, n (%)	161,443 (98.8)	96,985 (98.9)	64,458 (98.7)
Waste disposal place: fixed disposal, n (%)	155,458 (95.2)	93,406 (95.2)	62,052 (95)
Distance to drinking water source: shorter than median distance, n (%)	108,129 (66.2)	64,772 (66)	43,357 (66.4)
Vater filter available in HH, n (%)	19,271 (11.8)	11,628 (11.9)	7643 (11.7)

<sup>a</sup>Vi-TT: Vi-tetanus toxoid.

<sup>b</sup>BDT: Bangladeshi taka; currency conversion rate: US \$1 = BDT 110.48

<sup>c</sup>In total, 75 households had missing household expenditures, which were then substituted with the average household expenditure.

<sup>d</sup>WASH: waster, sanitation, and hygiene.

<sup>e</sup>HH: household.

#### Model Development, Training, and Validation

In the training subpopulation, having a private toilet (HR 0.57, 95% CI 0.43-0.77; P<.001); a safe source of drinking water, defined as a private tap, private well, bottled water, or water vendor (HR 0.65, 95% CI 0.49-0.87; P=.003); and availability

of a water filter in the household during observation (HR 0.73, 95% CI 0.45-1.18; *P*=.20) were associated with typhoid risk among members of the household (Table 2). These 3 variables were used to build a binary composite prediction rule defining "Better" *versus* "Not Better" household WASH using the recursive partitioning algorithm (Figure 2).



#### Tadesse et al

**Table 2.** Bivariate associations of binary WASH<sup>a</sup> variables with typhoid risk in the training subpopulation of the control clusters of the Vi-TT<sup>b</sup> cluster randomized trial, Dhaka, Bangladesh.

Variables	Yes					No				Hazard ratio (HR <sup>c</sup> ) <sup>d</sup>	
	n	Cases	PY <sup>e</sup>	IR <sup>f</sup> /100,000 PY	n	Cases	РҮ	IR/100,000 PY	HR (95% CI) <sup>g</sup>	<i>P</i> value <sup>g</sup>	
Toilet facility: pri- vate	41,155	86	57,718	149	56,932	174	67,381	258	0.57 (0.43- 0.77)	<.001	
Adult toilet: flush toilet	4617	11	6223	177	93,470	249	11,8876	209	1.07 (0.35- 3.30)	.91	
Child toilet: flush toilet	580	2	749	267	97,507	258	124,350	207	1.65 (0.41- 6.69)	.48	
Source of drinking water: private tap, well, or pump; wa- ter vendor	37,854	82	52,216	157	60,233	178	72,882	244	0.65 (0.49- 0.87)	.003	
Treated drinking water	73,527	183	92,350	198	24,560	77	32,748	235	0.91 (0.65- 1.27)	.57	
Freated cleaning water	300	1	420	238	97,787	259	124,678	208	1.13 (0.22- 5.78)	.89	
Hand wash before meals	73,019	191	92,584	206	25,068	69	32,514	212	1.15 (0.87- 1.52)	.33	
Hand wash after defecation	95,698	251	122,048	206	2389	9	3051	295	0.82 (0.48- 1.40)	.47	
Handwashing wa- er available in HH <sup>h</sup>	95,616	251	121,736	206	2424	9	3312	272	0.94 (0.33- 2.70)	.91	
Handwashing soap available in HH	96,985	257	123,590	208	1102	3	1508	199	1.3 (0.56- 3.03)	.537	
Waste disposal place: fixed dispos- al	93,406	240	118,789	202	4681	20	6309	317	0.77 (0.43- 1.37)	.37	
Distance to drink- ng water source: shorter than medi- an distance	64,772	175	83,175	210	33,315	85	41,923	203	1.11 (0.85- 1.45)	.46	
Water filter avail- able in HH	11,628	24	15,966	150	86,459	236	109,132	216	0.73 (0.45- 1.18)	.20	

<sup>a</sup>WASH: water, sanitation, and hygiene.

<sup>b</sup>Vi-TT: Vi-tetanus toxoid.

<sup>c</sup>HR: hazard ratio.

<sup>d</sup>Estimated from extended Cox proportional hazards model.

<sup>e</sup>PY: person years.

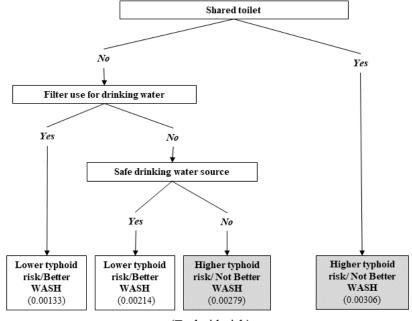
<sup>f</sup>IR: incidence rate.

<sup>g</sup>Calculated using robust SE assuming risk of typhoid is correlated within clusters.

<sup>h</sup>HH: household.

XSL•FO RenderX

**Figure 2.** Decision tree defining "Better" and "Not Better" household WASH using 3 binary WASH variables collected at first census. The figures within the parentheses in the final boxes of the decision tree indicate the proportion of individuals who tested positive for typhoid fever. Out of 40,329 households, 18 (0.0446%) experienced 2 typhoid episodes. Among these, 13 out of 28,230 (0.046%) were from households categorized as Not Better WASH, while 5 out of 12,099 (0.0413%) were from households categorized as Better WASH. WASH: water, sanitation, and hygiene.



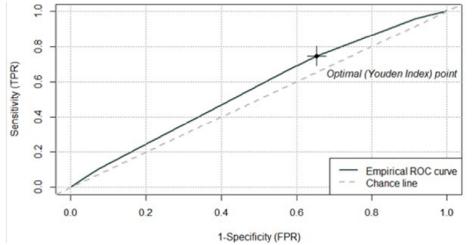
(Typhoid risk)

# Performance of the Composite WASH Variable Predicting Typhoid Fever

The decision tree is shown in Figure 2. The dominant bifurcation for the prediction of typhoid in the tree was having a private toilet. In the training subpopulation, an optimal cutoff value of 0.0025 maximized the Youden index using the ROC curve, with an AUC of 56% (95% CI 53-59; Figure 3). With this threshold,

the rule had a sensitivity of 74.6% (95% CI 68.9-80.0) and specificity of 34.7% (95% CI 34.4-35.0). Applying this rule to the validation subpopulation yielded comparable characteristics to those observed in the training subpopulation with a sensitivity and specificity of 74.9% (95% CI 67.8-81.1) and 35.2% (95% CI 34.9-35.6), respectively. The positive and negative predictive values were 0.3% (95% CI 0.26-0.35) and 99.8% (95% CI 99.7-99.9), respectively.

**Figure 3.** Receiver operator characteristic curve used to define the best cut off value to classify households as having "Better" versus "Not Better" WASH. AUC=56% (95% CI 53-59); best cutoff using the Youden index is 0.0025. AUC: area under the receiver operator characteristic curve; FPR: false positive rate; ROC: receiver operator characteristic; TPR, true positive rate; WASH: water, sanitation, and hygiene.



# Prediction of Typhoid Incidence by Household WASH Status in the Total Population

We next applied the WASH prediction rule to the overall population residing in the control clusters to predict typhoid risk. The incidence of typhoid fever in all age groups living in

https://publichealth.jmir.org/2023/1/e41207

RenderX

"Better" WASH households was 139/100,000 PYO compared to 252/100,000 PYO in those living in "Not Better" WASH households, a reduction in typhoid risk of 38% (95% CI 22-51; P<.001) after adjusting for covariates. The reduction in typhoid risk was largely attributed to the protection associated with living in "Better" WASH households among children <10 years

old (adjusted hazard ratio [aHR] 0.49, 95% CI 0.36-0.66; P<.001). Protective associations were aHR 0.35 (95% CI 0.20-0.61; P<.001) in children 2-4 years old, and aHR 0.58 (95% CI 0.36-0.92; P=.02) in 5-9 years old (Table 3 and Figure S1 in Multimedia Appendix 1).

The level of "Better" WASH cluster prevalence appeared inversely related to the log incidence of typhoid (Figure 4). However, the association between an increasing prevalence of "Better" WASH households in a cluster and typhoid risk was not significant in the adjusted multivariable model (aHR 0.996; 95% CI 0.986-1.006; P=.39) for each percent increase in the prevalence of "Better WASH" households in the cluster. The linear "Better" WASH cluster prevalence exhibited collinearity with household WASH status in the multivariable model, which precluded simultaneous estimation of the protective effects associated with living in "Better" WASH households and those associated with living in clusters with increasing prevalence of "Better" WASH households.

**Table 3.** Individual-level protection against typhoid associated with "Better" household  $WASH^a$  in the total population residing in the control clusters of the Vi-TT<sup>b</sup> cluster randomized trial, Dhaka, Bangladesh, stratified by age at onset of follow-up.

Years	"Better" WASH					"Not Better" WASH				Hazard ratio (HR <sup>c</sup> ) <sup>d</sup>			
	n <sup>e</sup>	Cases	$PY^{\mathrm{f}}$	IR <sup>g</sup> /100,000 PY	Ν	Cases	РҮ	IR/100,000 PY	Crude HR <sup>h</sup>	P value <sup>h</sup>	Adjusted HR <sup>h,i</sup>	P value <sup>h</sup>	
All	56,969	110	79,326	139	106,404	325	129,088	252	0.56 (0.44,0.71)	<.001	0.62 (0.49- 0.78)	<.001	
<2 years	3356	13	3883	335	6979	45	7211	624	0.54 (0.28,1.06)	.08	0.65 (0.32- 1.28)	.21	
2-4 years	3131	18	4424	407	6438	93	7919	1174	0.35 (0.22,0.57)	<.001	0.35 (0.20- 0.61)	<.001	
5-9 years	4894	26	7086	367	9819	84	12,586	667	0.57 (0.35,0.94)	.03	0.58 (0.36- 0.92)	.02	
10-14 years	4997	20	7278	275	9265	38	12,123	313	0.88 (0.5,1.56)	.66	0.92 (0.53- 1.59)	.77	
>15 years	40,591	33	56,655	58	73,903	65	89,250	73	0.76 (0.48,1.21)	.25	0.79 (0.50- 1.24)	.31	

<sup>a</sup>WASH: water, sanitation, and hygiene.

<sup>b</sup>Vi-TT: Vi-tetanus toxoid.

<sup>c</sup>HR: hazard ratio.

<sup>d</sup>Hazard ratios are estimated from extended Cox proportional hazards model.

<sup>e</sup>n: number; 100K: 100,000.

<sup>f</sup>PY: person years.

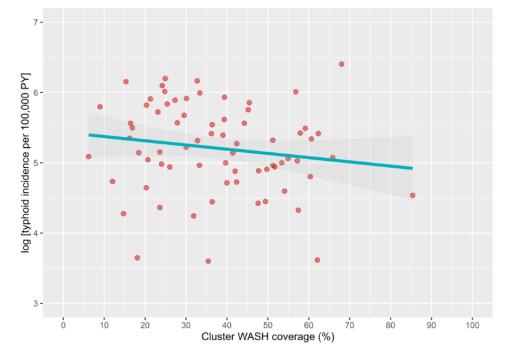
<sup>g</sup>IR: incidence rate.

<sup>h</sup>P values and CI are calculated using robust SE assuming risk of typhoid is correlated within clusters.

<sup>1</sup>Hazard ratio adjusted for the stratifying variables for randomization, including geographic ward, longer distance to study clinics than median distance, number of eligible children at baseline, and other baseline covariates age, Muslim religion; total monthly income; number of participants in a cluster.



**Figure 4.** Increasing levels of "Better" WASH cluster prevalence and the risk of typhoid in the total population residing in the control clusters of the Vi-TT cluster randomized trial. Out of 40,329 households, 28,230 (70%) were categorized as "Not Better" WASH households, and the remaining 12,099 (30%) fell under the "Better" WASH category. PY: person years; Vi-TT: Vi-tetanus toxoid; WASH: water, sanitation, and hygiene.



# Discussion

We developed a composite WASH variable for a poor urban population in Dhaka using a machine learning approach, where 3 binary WASH variables associated with typhoid risk were combined to discriminate "Better" versus "Not Better" household WASH status. The variable predicted the risk of typhoid fever after adjusting for confounding variables including household socioeconomic status, with individuals in households with "Better" WASH having a 38% lower risk of typhoid than those in households with "Not Better" WASH, a reduction in risk that appeared primarily attributable to reduced risk in children younger than the age of 10 years. The higher risk reduction in the younger age group with improved WASH practices can be potentially explained by the elevated incidence of typhoid among this demographic [23,24]. Further, there may be other behavioral disparities between adults and children, such as children primarily staying at home and adults being more frequently outdoors for work-related activities, which could contribute to this discrepancy.

The WASH variables defining the composite household WASH rule in this analysis are somewhat different from those found to predict the risk of typhoid in an earlier analysis in Kolkata [8], except for quality of drinking water, which contributed to the definition of the composite WASH variable in both analyses. These differences are, however, anticipated owing to the differences between the 2 settings including sociocultural characteristics, data collection time, and the different data instruments used for characterizing WASH in the 2 studies. Furthermore, both studies were primarily designed to evaluate typhoid vaccine protection, and neither study attempted in-depth characterization of WASH with the data collection instruments, so component criteria of the composite variables are very likely

RenderX

to be markers of WASH facilities and practices, rather than the specific facilities and practices per se that account for a reduced risk of typhoid. However, associations with typhoid in both settings were independent of socioeconomic status, underscoring that it was likely that better WASH rather than greater affluence was responsible for the protective associations with typhoid.

Some limitations should be noted while interpreting the results. First, as noted earlier, we used WASH variables collected as part of the randomized clinical trial of vaccines and were not ascertained for determination of the balance of the randomized intervention arms rather than as in a specific study of WASH. Second, our initial dichotomization of WASH variables in the population was based on perceived clinical and public health significance, although such judgements were made without knowledge of typhoid outcomes. Third, our creation of a composite WASH variable for households in a dichotomous fashion was done to create a simplified assessment of household level WASH and cluster prevalence of "Better" WASH, but this could have led to a potential loss of information on certain WASH variables, which might make our analysis conservative. An additional factor potentially influencing data validity is the Hawthorne effect, arising from alterations in observed behaviors.

Despite these limitations, our analysis has several strengths. First, typhoid incidence data used in this analysis were collected using a prospective cohort with comprehensive surveillance of typhoid fever with repeated censuses and WASH assessments that allowed analysis of a dynamic cohort in a highly mobile population. The intensified community engagement and proactive house-to-house visits by community health workers have played a role in ensuring that the proportion of participants across all age groups who refrained from seeking care for febrile episodes did not result in notable variations in care-seeking practices. Second, the detailed demarcation of the clusters using

natural structures, when possible, enabled unambiguous estimates of the "Better" WASH coverage in the clusters [6]. Third, external validation of the WASH prediction rule using a separate validation subpopulation revealed similar levels of sensitivity and specificity, suggesting that the prediction of typhoid by the composite variable was not exaggerated by overfitting the rule to the data used for derivation.

Simplified prediction rules, like the one presented in our study, offer great potential for various stakeholders involved in combating typhoid fever, including public health decision makers, researchers, and clinicians. These rules highlight the effectiveness of existing WASH adaptations, which can be implemented by most urban slum households to reduce susceptibility to typhoid in such environments. They can also be used to guide vaccine roll-out in high-risk populations and for community engagement initiatives aimed at increasing awareness of typhoid risk. Since the WASH data was collected from a diverse urban slum in Dhaka, the results can be extrapolated to other slums in Dhaka and similar regions with high typhoid burdens. Although further investigation is needed to address implementation barriers of "Better" WASH

parameters, well-designed risk communication strategies can facilitate the adoption of such adaptations [25]. While not directly related to the clinical diagnosis and treatment of individual typhoid cases, WASH interventions play a crucial role in the broader strategy by potentially reducing antibiotic resistance. By decreasing antibiotic usage, these interventions contribute to preserving the efficacy of antibiotics, thus expanding the range of treatment options available for combating typhoid fever.

In conclusion, our findings suggest that improvements in WASH that already exist in this slum population without specific external intervention are associated with a reduced risk of typhoid and lend optimism to the notion that simple and inexpensive improvements of WASH to protect against typhoid are possible. Research to develop and evaluate such interventions should be pursued. As well, it will be important to assess how better existing WASH interacted with and perhaps modified protection by Vi-TT typhoid vaccine tested in this CRT. We are carrying out these analyses and will report them in future publications.

# Acknowledgments

Funding for this study was from Bill and Melinda Gates Foundation (INV-025386).

# **Data Availability**

The data and code will be available upon reasonable request to the corresponding author.

# **Authors' Contributions**

JDC, FQ, FM, BTT, and FK conceptualized this study. Data analysis was done by FA, XL, JDC, BTT, FK, DRK, GP, JP, and MTI. Development of the first draft of this paper was by BTT, FK, MTI, XL, and JDC. JDC, FQ, FM, AIK, JHK, KZ, and DRK supervised for this paper. JI, SSYK, FC, TA, ABA, HJJ, and MH read, reviewed, and revised this paper. All the authors contributed to subsequent drafts. All the authors read and approved the final paper.

# **Conflicts of Interest**

None declared.

# Multimedia Appendix 1

Kaplan-Meier curve displaying time to onset of typhoid cases in Better versus Not Better WASH households. [PNG File , 83 KB-Multimedia Appendix 1]

# References

- GBD 2017 Typhoid and Paratyphoid Collaborators. The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Infect Dis 2019;19(4):369-381 [FREE Full text] [doi: 10.1016/S1473-3099(18)30685-6] [Medline: <u>30792131</u>]
- 2. Parry CM, Hien TT, Dougan G, White NJ, Farrar JJ. Typhoid fever. N Engl J Med 2002;347(22):1770-1782 [FREE Full text] [doi: 10.1056/NEJMra020201] [Medline: 12456854]
- 3. Typhoid. World Health Organization. Geneva; 2023. URL: <u>https://www.who.int/news-room/fact-sheets/detail/typhoid</u> [accessed 2023-11-07]
- 4. Marchello CS, Birkhold M, Crump JA. Complications and mortality of typhoid fever: a global systematic review and meta-analysis. J Infect 2020;81(6):902-910 [FREE Full text] [doi: 10.1016/j.jinf.2020.10.030] [Medline: 33144193]
- Fatima M, Kumar S, Hussain M, Memon NM, Vighio A, Syed MA, et al. Morbidity and mortality associated with typhoid fever among hospitalized patients in hyderabad district, Pakistan, 2017-2018: retrospective record review. JMIR Public Health Surveill 2021;7(5):e27268 [FREE Full text] [doi: 10.2196/27268] [Medline: 33999000]

RenderX

- Vighio A, Syed MA, Hussain I, Zia SM, Fatima M, Masood N, et al. Risk factors of extensively drug resistant typhoid fever among children in Karachi: case-control study. JMIR Public Health Surveill 2021;7(5):e27276 [FREE Full text] [doi: 10.2196/27276] [Medline: <u>33973861</u>]
- Stanaway JD, Atuhebwe PL, Luby SP, Crump JA. Assessing the feasibility of typhoid elimination. Clin Infect Dis 2020;71(Suppl 2):S179-S184 [FREE Full text] [doi: 10.1093/cid/ciaa585] [Medline: 32725226]
- 8. Im J, Islam MT, Ahmmed F, Kim DR, Khan AI, Zaman K, et al. Can existing improvements of water, sanitation, and hygiene (WASH) in urban slums reduce the burden of typhoid fever in these settings? Clin Infect Dis 2021;72(11):e720-e726 [FREE Full text] [doi: 10.1093/cid/ciaa1429] [Medline: 32964216]
- 9. Brockett S, Wolfe MK, Hamot A, Appiah GD, Mintz ED, Lantagne D. Associations among water, sanitation, and hygiene, and food exposures and typhoid fever in case-control studies: a systematic review and meta-analysis. Am J Trop Med Hyg 2020;103(3):1020-1031 [FREE Full text] [doi: 10.4269/ajtmh.19-0479] [Medline: 32700668]
- Breiman RF, Cosmas L, Njuguna H, Audi A, Olack B, Ochieng JB, et al. Population-based incidence of typhoid fever in an urban informal settlement and a rural area in Kenya: implications for typhoid vaccine use in Africa. PLoS One 2012;7(1):e29119 [FREE Full text] [doi: 10.1371/journal.pone.0029119] [Medline: 22276105]
- Mogasale VV, Ramani E, Mogasale V, Park JY, Wierzba TF. Estimating typhoid fever risk associated with lack of access to safe water: a systematic literature review. J Environ Public Health 2018;2018:9589208 [FREE Full text] [doi: 10.1155/2018/9589208] [Medline: 30174699]
- 12. Luby SP. Urban slums: a supportive ecosystem for typhoidal salmonellae. J Infect Dis 2018;218(suppl\_4):S250-S254 [FREE Full text] [doi: 10.1093/infdis/jiy324] [Medline: 30060082]
- 13. Akullian A, Ng'eno E, Matheson AI, Cosmas L, Macharia D, Fields B, et al. Environmental transmission of typhoid fever in an urban slum. PLoS Negl Trop Dis 2015;9(12):e0004212 [FREE Full text] [doi: 10.1371/journal.pntd.0004212] [Medline: 26633656]
- Naheed A, Ram PK, Brooks WA, Hossain MA, Parsons MB, Talukder KA, et al. Burden of typhoid and paratyphoid fever in a densely populated urban community, Dhaka, Bangladesh. Int J Infect Dis 2010;14(Suppl 3):e93-e99 [FREE Full text] [doi: 10.1016/j.ijid.2009.11.023] [Medline: 20236850]
- Theiss-Nyland K, Qadri F, Colin-Jones R, Zaman K, Khanam F, Liu X, et al. Assessing the impact of a Vi-polysaccharide conjugate vaccine in preventing typhoid infection among Bangladeshi children: a protocol for a phase IIIb trial. Clin Infect Dis 2019;68(Suppl 2):S74-S82 [FREE Full text] [doi: 10.1093/cid/ciy1107] [Medline: 30845333]
- Khanam F, Sheikh A, Sayeed MA, Bhuiyan MS, Choudhury FK, Salma U, et al. Evaluation of a typhoid/paratyphoid diagnostic assay (TPTest) detecting anti-Salmonella IgA in secretions of peripheral blood lymphocytes in patients in Dhaka, Bangladesh. PLoS Negl Trop Dis 2013;7(7):e2316 [FREE Full text] [doi: 10.1371/journal.pntd.0002316] [Medline: 23951368]
- Khanam F, Sayeed MA, Choudhury FK, Sheikh A, Ahmed D, Goswami D, et al. Typhoid fever in young children in Bangladesh: clinical findings, antibiotic susceptibility pattern and immune responses. PLoS Negl Trop Dis 2015;9(4):e0003619 [FREE Full text] [doi: 10.1371/journal.pntd.0003619] [Medline: 25849611]
- Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods 2009;14(4):323-348 [FREE Full text] [doi: 10.1037/a0016973] [Medline: 19968396]
- Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation 2007;115(5):654-657 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.105.594929] [Medline: 17283280]
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77 [FREE Full text] [doi: 10.1186/1471-2105-12-77] [Medline: 21414208]
- 21. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York, US: Springer; 2000.
- 22. Wickham H, François R, Henry L, Müller K, Vaughan D, Posit Software, PBC. dplyr: a grammar of data manipulation. R package version 0.7.6. 2018. URL: <u>https://cran.r-project.org/web/packages/dplyr/index.html</u> [accessed 2023-11-07]
- 23. Garrett DO, Longley AT, Aiemjoy K, Yousafzai MT, Hemlock C, Yu AT, et al. Incidence of typhoid and paratyphoid fever in Bangladesh, Nepal, and Pakistan: results of the Surveillance for enteric fever in Asia project. Lancet Glob Health 2022;10(7):e978-e988 [FREE Full text] [doi: 10.1016/S2214-109X(22)00119-X] [Medline: 35714648]
- 24. Marks F, von Kalckreuth V, Aaby P, Adu-Sarkodie Y, El Tayeb MA, Ali M, et al. Incidence of invasive salmonella disease in sub-Saharan Africa: a multicentre population-based surveillance study. Lancet Glob Health 2017;5(3):e310-e323 [FREE Full text] [doi: 10.1016/S2214-109X(17)30022-0] [Medline: 28193398]
- 25. Barac R, Als D, Radhakrishnan A, Gaffey MF, Bhutta ZA, Barwick M. Implementation of interventions for the control of typhoid fever in low- and middle-income countries. Am J Trop Med Hyg 2018;99(3\_Suppl):79-88 [FREE Full text] [doi: 10.4269/ajtmh.18-0110] [Medline: 30047365]

RenderX

#### Abbreviations

aHR: adjusted hazard ratio
AUC: area under the curve
CRT: cluster randomized trial
HR: hazard ratio
JE: Japanese encephalitis
PYO: person years of observation
ROC: receiver operating characteristic
TCV: typhoid conjugate vaccine
Vi-TT: Vi-tetanus toxoid
WASH: water, sanitation, and hygiene

Edited by A Mavragani, T Sanchez; submitted 18.07.22; peer-reviewed by B Padhi, S Bhattacharya, M Raimi, A Allam, W Klement, K Robb; comments to author 25.07.23; revised version received 30.08.23; accepted 05.10.23; published 20.11.23 <u>Please cite as:</u> Tadesse BT, Khanam F, Ahmmed F, Liu X, Islam MT, Kim DR, Kang SSY, Im J, Chowdhury F, Ahmed T, Aziz AB, Hoque M, Park J, Pak G, Jeon HJ, Zaman K, Khan AI, Kim JH, Marks F, Qadri F, Clemens JD Association Among Household Water, Sanitation, and Hygiene (WASH) Status and Typhoid Risk in Urban Slums: Prospective Cohort Study in Bangladesh JMIR Public Health Surveill 2023;9:e41207 URL: https://publichealth.jmir.org/2023/1/e41207 doi: 10.2196/41207 PMID:

©Birkneh Tilahun Tadesse, Farhana Khanam, Faisal Ahmmed, Xinxue Liu, Md Taufiqul Islam, Deok Ryun Kim, Sophie SY Kang, Justin Im, Fahima Chowdhury, Tasnuva Ahmed, Asma Binte Aziz, Masuma Hoque, Juyeon Park, Gideok Pak, Hyon Jin Jeon, Khalequ Zaman, Ashraful Islam Khan, Jerome H Kim, Florian Marks, Firdausi Qadri, John D Clemens. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.11.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

