Tsukuba Economics Working Papers
No. 2023-002

The benefit of clean water on child health: An empirical analysis with specific reference to *Escherichia Coli* water contamination

by

Ammazia Hanif
Yuko Nakano
Midori Matsushima

March 2023

UNIVERSITY OF TSUKUBA
Department of Economics
1-1-1 Tennodai
Tsukuba, Ibaraki 305-8571
JAPAN

Abstract

Microorganism-mediated degradation of water quality is a major public health concern in developing countries. Previous literature has shown an association between household water pollution and childhood diarrhoea; however, its effects on child growth and respiratory health have not been widely investigated. This study assessed the effects of household drinking water contaminated with *Escherichia coli (E. coli)* on child's weight-for-height and weight-for-age z-scores, acute respiratory infections (ARI), and diarrhoea incidence among five years children in Pakistan. We used district-level spatial information and the latest waves of unique Multiple Indicator Cluster Survey (MICS) data containing information on 'point-of-service delivery' (POS) and 'point-of-consumption' (POC) water quality, collected for the first time on a large scale in five regions of Pakistan. We employed an instrumental variable approach to address potential endogeneity issues in household drinking water quality, finding that POC drinking water contamination significantly affected children's weight-for-height and weight-for-age z-scores and ARI, in addition to its effects on diarrhoea. The sub-sample analyses indicated that the effects of contaminated water were particularly significant in children aged 6 months and older and in children who did not receive vitamin A supplements. To protect the children from growth failure and contracting ARI and diarrhoea, household water quality should be improved.

Keywords: Child Health, Undernutrition, ARI, E. coli, Water quality, Pakistan

Introduction

Ending undernutrition and reduction of infectious illness are critical prerequisites for sustainable development. Worldwide, around half of the under-five mortality (3 million) is attributable to undernutrition (Black et al., 2013), which is linked to amplified risk of morbidity and infections including respiratory illness (Rudan et al., 2013; Liu et al., 2015). Diarrhoea and acute respiratory infections (ARI) remain the leading causes of mortality and morbidity, especially in developing countries (Paulson et al., 2021; Osarogiagbon & Isara, 2018). Every year, from one billion childhood diarrhoea episodes, around 5 million result in death (Flückiger & Ludwig, 2022), and ARI contribute to 15% of deaths under the age of 5 worldwide. The cumulative health burden is evident because diarrhoea can impair child growth (Khalil et al.,2018) and trigger ARI (Walker, Perin, Katz, Tielsch, & Black, 2013)

Recent research and international development partners have focused on environmental contamination, which is a possible structural barricade to child development (Do, Joshi, & Stolper, 2018; Flückiger & Ludwig, 2022). Among environmental risks related to health, water pollution in particular, is considered an 'invisible threat' to the global development agenda, and contaminated water can significantly increase morbidity and affect labour productivity and well-being (Devoto, Duflo, Dupas, Parienté, & Pons, 2012). *E.coli* bacterial contamination[1] was found to be one of the main causes of failure to meet the global criteria for safely managing drinking water (Bain et al., 2021). Fecal exposure through poor drinking water, sanitation and hygiene is linked to approximately 62% diarrhoeal mortality, 13% ARI burden and 16% malnutrition cases (Prüss-Ustün et al., 2019).

---

[1] *Escherichia coli (E. coli)* - an indicator of faecal contamination is a species of faecal coliform group, easily cultured microorganisms and used to measure the water quality level.

Water pollution, malnutrition and infectious illness posing threat to the lives of children are severe in Pakistan. Bacterial contamination is increasing due to rapid urbanisation, water scarcity and poor sanitation. Access to safely managed drinking water is only 35.8% (WHO & UNICEF, 2022) showing that even if the source water is safe, unhygienic storage and water handling may lead to contamination (Julian, 2016). *E. coli* presence indicates water exposure to animal or human faeces (Barnes, Anderson, Mumma, Mahmud, & Cumming 2018; Santika et al., 2020; Usman, Gerber, & von Braun, 2019). As access to basic sanitation is 68.4% in Pakistan, inadequate sanitation can elevate the risk of microbial contamination due to extensive discharge of faecal sludge into water bodies (Ammazia, Yuko, & Midori, 2022 ; ; Zhang, 2012).

This study investigated the impact of E. coli contamination on weight-for-height and weight-for-age z-scores, ARI, and diarrhoea incidence in children up to the age of 5 years using Multiple Indicator Cluster Survey (MICS) data in Pakistan, which has one of the worst child health conditions in the world. MICS datasets have the unique feature of containing information on 'point of service delivery' (POS) and 'point-of-consumption' (POC) water quality, which was collected for the first time in a wide geographic area in Pakistan. This allows us to directly examine the impact of POC water quality on child health. Moreover, having information on POS enables us to identify community-level source water contamination. We use this community-level source water contamination as instrumental variable (IV) to account for the possible endogeneity of POC drinking water quality through our algorithm.

Our results document that a household using contaminated drinking water has a higher probability of diarrhoeal episodes by 7.9 percentage points and raises the ARI risk among children by 6.9 percentage points. Furthermore, our findings suggest that contaminated water is correlated with a decrease in children's weight-for-height by 0.31 and weight-for-age by 0.32 standard deviations. Based on these results, we infer that E. coli-induced episodes of diarrhoea could be a risk factor for undernutrition and ARI. Clean drinking water may improve respiratory health by preventing compromised immunity and micronutrient deficiencies. In addition to diarrhoea, environmental enteric dysfunction (EED), a subclinical gut disorder triggered by relentless faecal-oral contamination, is another possible route that associates water contamination with child growth (Modern et al., 2022; Campbell et al., 2018).

In addition, we conducted two subsample analyses that measured the differential effects of water quality on child health, focusing on children's age and micronutrient supplementation status. Dividing the sample by age group, we found that children aged 6 months and older were susceptible to E. coli contamination, possibly due to a larger intake of water and exposure to the contaminated environment than infants up to 6 months of age, who are generally exclusively breastfed. Moreover, periodic Vitamin A supplementation appears to protect children from contaminated disease environments as we found that children who were not given vitamin A supplement doses were more likely to be affected by poor water quality. These results suggest that vitamin A supplementation is an effective means of mitigating the negative effects of contaminated water on children's health and growth.

Our study makes two major contributions to this field. The first contribution has epidemiological importance as it reveals the impact of E. coli contamination of POC water on ARI and child growth. Although previous epidemiological studies have examined the impact of water quality on the incidence of diarrhoea (Brown, Proum, & Sobsey, 2008; Rahman, Kunwar, & Bohara, 2021; Usman et al., 2019; Khan, Hossain, Chakraborty, & Mistry, 2022; Hubbard et al., 2020; Ercumen et al., 2017; Fagerli et al., 2017; Luby et al., 2015; Gruber, Ercumen, & Jr, 2014; Goddard et al., 2020), height-for-age Z score (Goddard et al., 2020), self-reported maternal health (Kile et al., 2014), and child mortality (He & Perloff, 2016), the literature on the impact of E. coli contamination on ARI and child growth is scant. Since the evidence suggests that water, sanitation, and hygiene (WASH) interventions and the choice of water sources affect child growth and ARI (Ashraf et al., 2020; Huda et al., 2012; Luby et al., 2018; Null et al., 2018; Fathmawati, Rauf, & Indraswari, 2021; Shrestha, Vicendese, & Erbas, 2020; Swarthout et al., 2020), it is likely the case that E.coli contamination is causing negative impact on child growth and occurrence of ARI although some found no impacts (Swarthout et al., 2020). This is likely to be because of different nutritional status and different age of observation. The impact of E.coli contamination on child health, however, has not been well studied due to the lack of data on POC water quality. In addition, we investigate the differential impact of E.coli contamination on child health by age differences and Vitamin A intake. In doing so, this study provides more detailed policy implications than previous studies did.

The second contribution of our study relates to our data set and identification strategy. As discussed, MICS datasets have the unique feature of containing information on POS and POC water quality. This data enables us to directly investigate the effectiveness of POC contaminated water on child health by using community-level source water contamination as IV, since POC drinking water quality could be endogenously determined. This identification strategy provides robust results and adds evidence to the literature, as a limitation of previous studies is that they have largely ignored the endogeneity of household water quality. Although few exceptions are Rahman et al. (2021) and Usman et al. (2019), who also took endogeneity issues into consideration and found that microbial water contamination significantly affects the incidence of child diarrhoea, none of the other previous studies have done so.

The reminder of this paper is organised as follows. In Section 2, we explain the datasets. Section 3 outlines the empirical strategy. Section 4 documents the regression results of the analyses across several specifications, and Section 5 discusses the results. Finally, Section 6 concludes the paper.

Data

Sampling

This study utilised data from the latest round of the Multiple Indicator Cluster Surveys (MICS) in Pakistan; a provincially representative cross-sectional survey followed a two-stage stratified random sampling technique. We used the MICS survey waves from five regions in the country from 2017 to 2021[2]. During the survey, the first primary enumeration areas were selected from rural and urban areas within each district. Subsequently, a systematic sample of 20 households was randomly selected from each enumeration area (hereafter termed 'cluster'). A unique feature of MICS is the addition of a new water quality module developed by MICS and the Joint Monitoring Program (JMP) (*UNICEF MICS*, 2018) which was released as a part of latest round MICS6 .[3] Worldwide, few datasets include bacteriological water quality indicators at a national scale owing to collection challenges, and in most cases, only the quality of piped water is measured. In our datasets, water quality data at the POC and POS were collected separately for the first time on a large scale in Pakistan using a universally accepted method (Khan et al., 2017).

---

[2] Five survey waves include provinces MICS Punjab 2017–18, MICS Sindh 2018–19, MICS KPK 2019, MICS Baluchistan 2019–20 and a region MICS AJK 2020–21.

[3] Methods were developed for direct water testing in MICS. Generally, laboratories conduct water tests, which often suffer from logistic challenges in term of transportation and time frame.

In the MICS datasets, 113,233 households were interviewed with information on child anthropometry and other health outcomes from study regions of the country. For water quality testing, 18,493 households were randomly selected. We limited our sample where households gave consent for the water test; enumerators visited the source for E. coli test, and information for both the source and household water test was available. For the analyses, 12,671 observations remained for children aged 0–60 months after cleaning the data.

Outcome variables

This study used four outcome variables. First, we used diarrhoea prevalence, which took the value of 1 if the child's mother/caregiver reported the child had had diarrhoea 2 weeks before the survey. Second, acute respiratory infections (ARI) symptoms were used, which took the value of 1 if the child had experienced rapid or difficult breathing along with a persistent cough and the symptoms were associated with chest problems and a blocked nose 2 weeks before the survey4. The third and fourth outcomes were child growth: weight-for-height and weight-for-age z-scores, which are the standard deviations of the anthropometry of the child based on the growth standards of the WHO. Weight-for-age of less than -2 represents underweight, while a child's weight-for-height of less than -2 signifies wasting – a short-term undernutrition indicator owing to acute weight loss.

Main explanatory variables

The key explanatory variable of POC-contaminated drinking water was a dummy variable, which took 1 if E. coli was found in a household's POC water[5]. The WHO drinking water

---

[4] Respiratory rate is a valuable clinical sign for diagnosing acute lower respiratory infections (e.g. pneumonia and bronchiolitis) in children coughing and breathing rapidly.

[5] The following precise question was asked: Could you please provide me with a glass of water that members of your households usually drink?

quality guidelines recommend that E. coli must not be detectable in 100-ml water samples (WHO, 2022. A water test was performed on the samples provided by the households within 30 minutes of collection followed by 24–48 hours of incubation. E. coli colonies were counted as colony-forming units (CFU) /100 ml of water sample. Since each household use water from different source, the source water sample was also collected for each household, and the E. coli contamination was calculated by using similar method used for household water test.

Fig. 1 shows the specific E. coli count (CFU/100 ml, ranging from 0 to 100 and above) at the POC in the left panel and at the source in the right panel. In our sample, drinking water of around 78% households was found contaminated with E. coli, and the percentage of households having contaminated water at the source was around 59%. The variation in E. coli concentration at the POC and source showed that a higher E. coli risk (100 or above CFU) was prevalent at POC than at POS, suggesting that contamination occurs during the process of fetching and handling drinking water.

[Fig. 1]

In the regression analysis, we controlled for other children, mothers and household characteristics. Child-level variables included gender, age and prenatal check-ups. The mother and household characteristics comprise the age and education of the mother and household head, the household head's gender, maternal smoking, household size, number of children up to 5 years of age, cooking place, number of household members, household assets, access to flush toilets as indicator of improved sanitation[6], livestock ownership and

---

[6] Flush toilets reduced child weight loss, offer non-health welfare (e.g. time saving) and augments satisfaction (Wang & Shen, 2022; Rahman et al., 2021).

survey year dummy. The community-level variables included rural and urban dummies. Appendix Table 1 presents the summary statistics for the variables used in the analyses.

District characteristics

We supplemented the information from the MICS with additional geospatial datasets to account for district-level characteristics that may affect health and water quality. Using raster datasets from QGIS, we constructed district-level data on rainfall, population density, slope and temperature. Based on the district's geocoded position, we link the MICS data to information on district-level characteristics. For rainfall data, we used the Climate Hazards Group InfraRed Precipitation with Station data 2.0 (CHIRPS) (Funk et al., 2015). For population density, we used WorldPop data (WorldPop; Bondarenko & Maksym, 2020). The district slope was calculated using ALOS DSM: Global 30m v3.2 data provided by the JAXA Earth Observation Research Centre (EORC). For temperature, we used the dataset provided by the NASA LP DAAC at the USGS EROS Centre. To control for the economic status in each district, we used 3-arc-second (approximately 100 m at the equator) night-time light data constructed by NOAA's National Centres for Environmental Information as a proxy. In addition, we use malaria prevalence- a serious public health issue, as a proxy for malaria risk (*Plasmodium vivax* parasite rate in general population). To obtain information on the average malaria prevalence rate, Malaria-Atlas map was used. We overlayed the map estimations on each district location.

In Fig.2, using district-level spatial information, we mapped child health data from the sample and plotted the district-level risk prevalence of E. coli at POS and POC level and health-related outcomes. The figure shows that diarrhoea and ARI prevalence were higher in areas with a high possibility of E. coli concentrations in POC water. A similar tendency was observed for child growth indicators.

[Fig 2]

Empirical strategy

This section describes the empirical strategy employed to estimate the impact of household POC water quality on children's health. A linear probability model (LPM) and two-stage least squares (2SLS) were used. Following Rahman et al. (2021), we use Grossman's (1972) health production function approach as the motivation for our empirical estimation. Following the conceptual framework of this approach, we estimate the model below, where the health of child *i* is a function of the POC drinking water quality level and other demographic and socioeconomic factors:

$$\mathrm{Y}_{ijkd} = \beta_0 + \beta_1 \mathrm{WQ}_{jkd} + \beta_2 \mathrm{H}_{ijkd} + \delta_d + \theta_{dr} + \theta_t + e_{ijkd} \qquad [1]$$

where Y_{ijkd} refers to the health of child *i* of household *j* in cluster *k* of district *d*. Diarrhoea prevalence, weight-for-height, weight-for-age z-score and ARI are the indicators of child health and the outcome variables of interest. WQ_{jkd} is a dichotomous variable that denotes POC water quality (presence of E. coli) in household *j* in cluster *k*. H_{ijkd} is a vector for control variables at the household and individual levels. δ_d controls district-level characteristics including population density, average rainfall and temperature during the survey year, average night-time light and slope of each district. θ_{dr} controls for district-times-rural fixed effects. Because rural and urban areas have different characteristics within each district in terms of socioeconomic and health status, we include district-times-rural fixed effects to account for any region-specific variability. θ_t is the time-fixed effect, while e_{ijkd} is the error term that explains the variation of Y_{ijkd}. We clustered standard errors at the MICS cluster level.

Since each household can decide whether to take any means to improve water quality, our model may suffer from a possible endogeneity problem. For instance, some households may have unobserved perceptions, preferences, behaviours, and awareness that allow them to improve water quality and children's health. If households are keen on health, for example, they could invest not only in improving water quality but also in other measures to benefit their children's health, which may cause bias in our estimates. To circumvent such issues, we use 2SLS estimation and estimate the following regression equation as the first-stage equation of the 2SLS estimation of equation 1:

$$\mathrm{WQ}_{jkd} = \gamma_0 + \gamma_1 \mathrm{SWQ}_{kd} + \gamma_4 \mathrm{H}_{ijkd} + \delta_d + \theta_{dr} + \theta_t + v_{jkd} \qquad [2]$$

where the excluded instrumental variable is community-level POS water contamination, indicated as SWQ_{kd}. As we discussed, in the MICS, the POS water quality data were measured at the household level because each household used water from different sources. As each household cannot change the quality of the water at the source, the POS water quality should be largely exogenous for each household. There is, however, some possibility of endogeneity because each household may be able to choose sources to take their water. To circumvent this problem, we used non- self community level POS water contamination by measuring the leave-out means[7]. If the community-level source contamination increases, it would positively affect household water contamination. In addition, because the average POS water contamination most likely affects children's health mainly through drinking water quality, the variable can satisfy the exclusion restriction, although there is no direct way to test it. Therefore, we considered community-level POS water contamination as a plausible instrument for our study.

Results

Descriptive statistics

In Table 1, we compare child health and growth outcomes among households with POC drinking water with and without E. coli contamination. Table 1 shows that diarrhoea prevalence is higher by 4.8% among children from households with contaminated drinking water than households with E. coli-free water. The average weight-for-height and weight-for-age z-scores were lower among children from households exposed to E. coli contamination than among those using uncontaminated water. On average, the weight-for-height is -0.42 for children exposed to contaminated water and -0.32 for those who drink water free of E. coli. We can observe a similar tendency for weight-for-age. The weight-for-age of children with contaminated water was -1.32, while that of children with uncontaminated water was -1.10, and the difference was statistically significant. Further, the ARI prevalence is also higher by 1.3% among children exposed to contaminated water than among households using uncontaminated water, although the difference is not statistically significant.

[Table 1]

Fig 3 also shows dependence among E. coli and child health indicators. The left section presents the positive association, on average, between diarrhoea and ARI prevalence and the E. coli presence in POC water. Which indicates better child health association with the absence of E. coli in drinking water. When we consider child growth indicators, we observe a negative association, on average, between weight-for-height and weight-for-age z-scores and contaminated drinking water in the right section of Fig. 3. Which shows that improved child growth was associated with the E. coli absence in drinking water.

[Fig.3]

Main results

Estimates of the impact of POC E. coli in water on child diarrhoea, weight-for-height, weight-for-age, and ARI prevalence are presented in Table 2. The estimated coefficients of contaminated water on weight-for-height and weight-for-age were negative and statistically significant. The results show that the having E. coli in POC water decreases the child's weight-for-height and weight-for-age z-score by 0.067 and 0.076, respectively. In contrast, the estimated coefficients from OLS were insignificant for diarrhoea and ARI prevalence when we included district-times-rural fixed effects and time-fixed effects in the model. Conversely, the estimated coefficients of contaminated water for weight-for-height and weight-for-age were negative and statistically significant.

[Table 2]

To circumvent this problem, we estimated the same model using community-level POS water quality as IV for household-level POC water quality. The reduced- form estimates in Appendix Table 2 suggest that households in a cluster with contaminated water have a higher probability of diarrhoea incidence by 1.9 percentage points and of ARI by 1.6 percentage points than those having an uncontaminated water source. The estimated coefficients of the weight-for-height and weight-for-age z-scores are -0.074 and -0.077, respectively, and are statistically significant.

Panel A of Table 3 shows the first-stage estimation results. The dependent variable is POC water quality, while IV is community-level POS water contamination. The results show a strong positive correlation between POC water quality and IV, and diagnostic tests provide evidence of the validity of IV. For all outcomes, the null hypothesis for under-identification is rejected based on the LM version of the Kleibergen-Paap rk statistic.

In Panel B of Table 3, we estimate β_1 by applying 2SLS that captures the impact of POC water quality on child health outcomes. Column 1 shows that the probability of diarrhoea in children increases by 7.9 percentage points in households with contaminated drinking water compared to those using uncontaminated water, this translates to 44 % increase in diarrhoea risk. The results for weight-for-height and weight-for-age as alternative outcomes are also shown in columns 2 and 3. Consistent with childhood diarrhoea triggering weight loss, particularly in the short run, we found that POC-contaminated drinking water was associated with a decrease in children's weight-for-height by 0.31 standard deviations and their weight-for-age by 0.32 standard deviations. Column 4 shows that the probability of ARI risk among children increased by 6.9 percentage points in households with microbial-contaminated water. The IV estimates of the effect of water contamination on child health are larger than those from OLS regressions and are all statistically significant. Although some coefficients of the OLS estimates are insignificant, the direction of the impact is the same as that of the IV estimates, suggesting the robustness of our results.

[Table 3]

Differential effects of E. coli contamination

In the previous estimation, we assumed that the effects of E. coli contamination are the same across all sample children. We further assessed whether the effects of POC water quality varied across the different age groups as E. coli contamination may not similarly affect all children. The results presented in Table 4 show stark differences in the impact of contaminated water on child health among different age groups. The IV results in columns 1–4 show the impact of water quality on the health of children aged 0–5 months, while columns 5–8 show such impacts on children 6–60 months of age. The estimated coefficients in columns 5-8 indicate that the link between E. coli-contaminated water and the risk of diarrhoea and ARI incidence was statistically significant among children older than 6 months, with an average effect of 9.4 and 7.7 percentage points respectively. Moreover, contaminated water was significantly associated with a 0.34 and 0.35 standard deviation decrease in weight-for-height and weight-for-age z-scores in children older than 6 months. In contrast, the results in columns 1–4 show that water pollution does not affect children of less than 6 months of age. This is plausible because the introduction of complementary feeding and partial weaning from the age of 6 months increases a child's exposure to a contaminated environment[8]. Children exclusively breastfed during the initial 6 months may be protected naturally from pathogen exposure through contaminated food and water, having less ARI, diarrhoeal morbidity and mortality (Mulatu et al., 2021; Arifeen et al., 2001; VanDerslice et al., 1994;) and growth failure (Dharel et al., 2020). Although the results are not shown, we found that exclusively breastfed children (aged 0–5 months) were not significantly affected by water contamination.

[Table 4]

In addition, we analysed whether the intake of vitamin A supplements mitigated the negative effects of E. coli-induced risk of diarrhoea and ARI. Vitamin A deficiency (VAD) remains a public health concern across the globe including in Pakistan, where its frequency is more than 51%. According to previous studies, vitamin A (anti-infectious) is significant for gastrointestinal and respiratory epithelial regeneration and reduces immune disorders, mortality, and morbidity among children aged 6–59 months (Kanakala et al., 2019; Imdad et al., 2011; Barreto et al., 1994; WHO, 2013). Periodic vitamin A supplementation could have the capacity to protect children by boosting their immune systems from contaminated disease environments and related health risks (WHO, 2022) .

---

[8] In terms of food intake for the previous day, the information was only solicited for children of up to 2 years at the time of the survey.

To examine these hypotheses, we conducted subsample analyses of the impact of contaminated water on child health and growth for those who were supplemented with vitamin A and those who were not. In our analyses, the vitamin A variable took the value of 1 if the child aged 6–60 months had been provided with vitamin A supplementation in the previous 6 months. Table 5 shows the differential effects of contaminated water on the health and growth of children with and without vitamin A supplementation. Columns 1–4 show the effects of E. coli on the outcome variables for the vitamin A-supplemented group, while columns 5–8 show the effects on the non-supplemented group. The results in columns 5 and 8 show that contaminated water was statistically significantly associated with the risk of diarrhoea and ARI incidence, with an effect of 22 and 15 percentage points respectively, in the non-vitamin A-supplemented group. Likewise, as shown in column 6, contaminated water was significantly associated with a 0.56 standard deviation decrease in children's weight-for-height in the non-vitamin A-supplemented group. In contrast, we did not observe any significant effects of E. coli contamination on health among children supplemented with vitamin A, except for a negative coefficient for weight-for-age, for which we do not have a clear explanation. These results provide suggestive evidence to support our hypothesis that micronutrient supplementation can mitigate E. coli-induced health risks.

[Table 5]

Robustness checks

We conducted an array of robustness checks to document the validity and stability of our results. First, in Appendix Table 3, we examined whether the estimated 2SLS coefficients remained robust after excluding extensive covariates. In the main models, we controlled for variables related to the child, the mother, the household head, the household, district-level characteristics, district-times-rural fixed effects and time-fixed effects. We gradually excluded district and household controls in columns 1, 2, 4, 5, 7, 8, 10 and 11, while we included all these variables in columns 3, 6, 9 and 12 (the same as the main results shown in Table 2). The results show that the coefficient of water quality remained largely steady despite the inclusion or exclusion of various covariates.

Second, we added several variables which could affect water quality in the first- and second-stage regressions of our IV estimation. We included the means to extract water (motor, hand pump, and tube-well dummy), piped water dummy, water treatment status dummy, handwashing dummy, flush toilet dummy and animal ownership. As these can potentially determine water quality, excluding them may cause an omitted variable bias, especially in the first-stage estimation. At the same time, however, these variables are endogenous, and thus, we excluded them from our main analyses. Appendix Table 4 shows the first-stage and IV estimates with and without these variables. Column 1 in Panel A show first-stage results of Table 2 and column 2 controls for the effect of potential water quality determinants. Columns 1, 3, 5 and 7 in Panel B show the 2SLS results of Table 2, without including variables that can determine water quality. In columns 2, 4, 6, and 8 of Panel B, we include potential water quality determinants. The results show that the coefficients are almost the same for the estimation results with and without these covariates, suggesting the robustness of the results.

Lastly, in Appendix Table 5, as a sensitivity analysis, we estimated our models by measuring the E. coli coliform as a continuous variable. Columns 1 and 4 show that a 1% increase in E. coli (CFU/100 ml) in water increased the probability of diarrhoea by 0.014 percentage points and ARI by 0.013 percentage points. Such an increase in E. coli was associated with a decrease in children’s weight-for-height by 0.058 standard deviations and in their weight-for-age by 0.059 standard deviations; these effects were statistically significant. Overall, the robustness tests suggested that the estimated findings are robust and consistent.

Discussion

This study investigated the impact of household water quality on the incidence of childhood diarrhoea, weight-for-height, weight-for-age and ARI using MICS data in Pakistan. Globally, only a few datasets include bacteriological water quality indicators at the national scale. MICS Pakistan is a unique data set that allows us to directly examine the impact of POC water quality on children's growth and ARI. Moreover, previous studies have largely ignored the endogeneity of household drinking water quality, which might have caused bias in their estimates. We addressed these concerns by employing the two-stage least squares method and found that poor drinking water quality significantly and negatively affected child growth, ARI, and increased the incidence of diarrhoea. In addition, we examined the differential impact of water quality on child health by age and vitamin A supplementation by conducting sub-sample analyses.

Our findings suggest that POC water contamination elevates the risk of malnutrition, diarrhoea, and ARI. Negative impact of E. coli contamination on diarrhoea is consistent with previous studies (e.g. Khan et al., 2022; Rahman et al., 2021; Usman et al., 2019; Luby et al., 2015), and we added evidence to the growing literature on negative impact of E. coli contamination on weight-for-height, weight-for-age, and ARI by employing rigorous empirical estimation.

Age-based sub-sample analyses revealed that contaminated water significantly affects child health indicators, particularly when a child turns 6 months or more. A reason for this could be that after 6 months, children consume large quantities of water and are at a higher risk of being infected through contact with contaminated water. It is possible that infants (0–5 months) are naturally protected by breastfeeding; thus, we did not observe a significant negative impact of contaminated water on their health status. We also examined the differential impacts of water quality on child health depending on the provision of micronutrient supplements to children aged 6–60 months. Our results suggest that bacterial contamination did not have a significant effect on the incidence of diarrhoea, ARI, and weight-for-height z-score in children who were given vitamin A supplementation in the previous 6 months, while their counterparts were significantly affected. These results imply the effectiveness of vitamin A supplementation in mitigating the negative impacts of contaminated water on children's health.

Despite our contributions, this study has two main limitations. First, water quality was measured at a single point in time; therefore, we could only rely on cross-sectional variations to estimate the health impact of water quality. Undernutrition reflects the accumulated effects of nutritional deficiency; therefore, a longitudinal study on the effects of water quality on chronic undernutrition in children should be conducted in the future. Second, we did not have information on the seasonal variability of water quality and E. coli levels; thus, we could not examine whether rainy or dry seasons affect water contamination and child health differently.

Conclusion

Pakistan is one of the countries that is most susceptible to respiratory infections and diarrhoea-related mortality, and malnutrition in South Asia. This study examined the impact of microbial water contamination on childhood diarrhoea, weight-for-height, weight-for-age, and ARI prevalence using data from the latest waves of MICS in five regions of Pakistan during 2017–2021. Our key findings indicate that the presence of E. coli in POC water, which is generally used for drinking, significantly decreases a child's weight-for-height and weight-for-age and increases the risk of ARI and diarrhoea.

From our findings, we can highlight certain policy recommendations. First, financial allocations should be diverted towards the provision of safely managed drinking water to improve children's health. For example, the establishment of waste-management treatment plants and sewerage schemes may contribute to improving children's health by reducing water contamination. Second, families with children should realise the importance of a clean environment for their growing children and that micronutrient supplementation and exclusive breastfeeding should be promoted. In Pakistan, currently, only approximately 37.7% of women exercise exclusive breastfeeding, and more than 50% of children suffer from vitamin A deficiency (National Nutrition Survey, 2018). The promotion of breastfeeding and supplementation with vitamin A would not only have a direct impact on child health but also indirectly, by mitigating the negative impact of drinking water contamination.

Considering that the issue of poor child health is getting severe possibly due to recent massive floods in the country negatively affecting water safety and sanitation systems (UNICEF, 2022), it is necessary for us to conduct longitudinal analyses and observe long-term impacts of water contamination as well as how infrastructure and hygiene practice can protect child health.

Acknowledgments

We would like to acknowledge the contributions of UNICEF in providing open access to the MICS datasets. We are grateful for constructive comments from Hisahiro Naito, Fatwa Ramdani, Kenichi Kashiwagi, Mari Minowa, Yu Zhengfei, Moges Abu Girma, Mohammad Abdul Malek, Taiyo Fukai and Yoshinori Kurokawa.

References

Arifeen, S., Black, R. E., Antelman, G., Baqui, A., Caulfield, L., & Becker, S. (2001). Exclusive breastfeeding reduces acute respiratory infection and diarrhea deaths among infants in Dhaka slums. *Pediatrics*, *108*(4), E67. https://doi.org/10.1542/peds.108.4.e67

Ashraf, S., Islam, M., Unicomb, L., Rahman, M., Winch, P. J., Arnold, B. F., Benjamin-Chung, J., Ram, P. K., Colford, J. M., & Luby, S. P. (2020). Effect of improved water quality, sanitation, hygiene, and nutrition interventions on respiratory illness in young children in rural Bangladesh: a multi-arm cluster-randomized controlled trial. *The American Journal of Tropical Medicine and Hygiene*, *102*(5), 1124–1130. https://doi.org/10.4269/ajtmh.19-0769

Bain, R., Johnston, R., Khan, S., Hancioglu, A., & Slaymaker, T. (2021). Monitoring drinking water quality in nationally representative household surveys in low- and middle-income countries: cross-sectional analysis of 27 multiple indicator cluster surveys 2014–2020. *Environmental Health Perspectives*, *129*(9), 097010. https://doi.org/10.1289/EHP8459

Barnes, A. N., Anderson, J. D., Mumma, J., Mahmud, Z. H., & Cumming, O. (2018). The association between domestic animal presence and ownership and household drinking water contamination among peri-urban communities of Kisumu, Kenya. *PloS One*, *13*(6), e0197587. https://doi.org/10.1371/journal.pone.0197587

Barreto, M. L., Santos, L. M., Assis, A. M., Araújo, M. P., Farenzena, G. G., Santos, P. A., & Fiaccone, R. L. (1994). Effect of vitamin A supplementation on diarrhoea and acute lower-respiratory-tract infections in young children in Brazil. *Lancet (London, England)*, *344*(8917), 228–231. https://doi.org/10.1016/s0140-6736(94)92998-x

Bashir, S., Aslam, Z., Niazi, N. K., Khan, M. I., & Chen, Z. (2021). Impacts of water quality on human health in Pakistan. In M. A. Watto, M. Mitchell, & S. Bashir (Eds.), *Water Resources of Pakistan: Issues and Impacts* (pp. 225–247). Springer International Publishing. https://doi.org/10.1007/978-3-030-65679-9_12

Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., de Onis, M., Ezzati, M., Grantham-McGregor, S., Katz, J., Martorell, R., & Uauy, R. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, *382*(9890), 427–451. https://doi.org/10.1016/S0140-6736(13)60937-X

Brown, J. M., Proum, S., & Sobsey, M. D. (2008). Escherichia coli in household drinking water and diarrheal disease risk: Evidence from Cambodia. *Water Science and Technology*, *58*(4), 757–763. https://doi.org/10.2166/wst.2008.439

Campbell, R. K., Schulze, K. J., Shaikh, S., Raqib, R., Wu, L. S. F., Ali, H., Mehra, S., West, K. P., & Christian, P. (2018). Environmental enteric dysfunction and systemic inflammation predict reduced weight but not length gain in rural Bangladeshi children. *British Journal of Nutrition*, *119*(4), 407–414. https://doi.org/10.1017/S0007114517003683

Devoto, F., Duflo, E., Dupas, P., Parienté, W., & Pons, V. (2012). Happiness on Tap: Piped water adoption in urban Morocco. *American Economic Journal: Economic Policy*, *4*(4), 68–99. https://doi.org/10.1257/pol.4.4.68

Do, Quy-Toan, Shareen Joshi, and Samuel Stolper. 2018. “Can Environmental Policy Reduce Infant Mortality? Evidence from the Ganga Pollution Cases.” *Journal of Development Economics 133*:306–25. doi: 10.1016/j.jdeveco.2018.03.001.

Dharel, D., Dhungana, R., Basnet, S., Gautam, S., Dhungana, A., Dudani, R., & Bhattarai, A. (2020). Breastfeeding practices within the first six months of age in mid-western and eastern regions of Nepal: A health facility-based cross-sectional study. *BMC Pregnancy and Childbirth*, *20*(1), 59. https://doi.org/10.1186/s12884-020-2754-0

Ercumen, A., Arnold, B. F., Naser, A. Mohd., Unicomb, L., Colford Jr, J. M., & Luby, S. P. (2017). Potential sources of bias in the use of Escherichia coli to measure waterborne diarrhoea risk in low-income settings. *Tropical Medicine & International Health*, *22*(1), 2–11. https://doi.org/10.1111/tmi.12803

Ercumen, A., Naser, A. M., Unicomb, L., Arnold, B. F., Jr, J. M. C., & Luby, S. P. (2015). Effects of source- versus household contamination of tubewell water on child diarrhea in rural Bangladesh: A randomized controlled trial. *PLOS ONE*, *10*(3), e0121907. https://doi.org/10.1371/journal.pone.0121907

Fagerli, K., Trivedi, K. K., Sodha, S. V., Blanton, E., Ati, A., Nguyen, T., Delea, K. C., Ainslie, R., Figueroa, M. E., Kim, S., & Quick, R. (2017). Comparison of boiling and chlorination on the quality of stored drinking water and childhood diarrhoea in Indonesian households. *Epidemiology and Infection*, *145*(15), 3294–3302. https://doi.org/10.1017/S0950268817002217

Fathmawati, F., Rauf, S., & Indraswari, B. W. (2021). Factors related with the incidence of acute respiratory infections in toddlers in Sleman, Yogyakarta, Indonesia: Evidence from the Sleman Health and Demographic Surveillance System. *PLOS ONE*, *16*(9), e0257881. https://doi.org/10.1371/journal.pone.0257881

Flückiger, M., & Ludwig, M. (2022). Temperature and risk of diarrhoea among children in Sub-Saharan Africa. *World Development*, *160*, 106070. https://doi.org/10.1016/j.worlddev.2022.106070

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data*, *2*(1), Article 1. https://doi.org/10.1038/sdata.2015.66

Getachew, A., Tadie, A., G.Hiwot, M., Guadu, T., Haile, D., G.Cherkos, T., Gizaw, Z., & Alemayehu, M. (2018). Environmental factors of diarrhea prevalence among under five children in rural area of North Gondar zone, Ethiopia. *Italian Journal of Pediatrics*, *44*(1), 95. https://doi.org/10.1186/s13052-018-0540-7

Goddard, F. G. B., Pickering, A. J., Ercumen, A., Brown, J., Chang, H. H., & Clasen, T. (2020). Faecal contamination of the environment and child health: A systematic review and individual participant data meta-analysis. *The Lancet. Planetary Health*, *4*(9), e405–e415. https://doi.org/10.1016/S2542-5196(20)30195-9

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*. https://doi.org/10.1086/259880

Gruber, J. S., Ercumen, A., & Jr, J. M. C. (2014). Coliform bacteria as indicators of diarrheal risk in household drinking water: Systematic review and reta-analysis. *PLOS ONE*, *9*(9), e107429. https://doi.org/10.1371/journal.pone.0107429

Hanif, A., Nakano Y., & Matsushima M. (2022). *The impact of access to improved sanitation facilities on child health in Pakistan*. Tsukuba Economics Working Papers No. 2022-003. https://pepp.hass.tsukuba.ac.jp/RePEc/2022-003.pd

He, G., & Perloff, J. M. (2016). Surface water quality and infant mortality in China. *Economic Development and Cultural Change*, *65*(1), 119–139. https://doi.org/10.1086/687603

Hubbard, S. C., Meltzer, M. I., Kim, S., Malambo, W., Thornton, A. T., Shankar, M. B., Adhikari, B. B., Jeon, S., Bampoe, V. D., Cunningham, L. C., Murphy, J. L., Derado, G., Mintz, E. D., Mwale, F. K., Chizema-Kawesha, E., & Brunkard, J. M. (2020). Household illness and associated water and sanitation factors in peri-urban Lusaka, Zambia, 2016–2017. *Npj Clean Water*, *3*(1), Article 1. https://doi.org/10.1038/s41545-020-0076-4

Huda, T. M. N., Unicomb, L., Johnston, R. B., Halder, A. K., Yushuf Sharker, M. A., & Luby, S. P. (2012). Interim evaluation of a large-scale sanitation, hygiene and water improvement programme on childhood diarrhea and respiratory disease in rural Bangladesh. *Social Science & Medicine (1982)*, *75*(4), 604–611. https://doi.org/10.1016/j.socscimed.2011.10.042

Imdad, A., Yakoob, M. Y., Sudfeld, C., Haider, B. A., Black, R. E., & Bhutta, Z. A. (2011). Impact of vitamin A supplementation on infant and childhood mortality. *BMC Public Health*, 11(Suppl 3), S20. https://doi.org/10.1186/1471-2458-11-S3-S20

Julian, T. R. (2016). Environmental transmission of diarrheal pathogens in low- and middle-income countries. *Environmental Science: Processes & Impacts*, *18*(8), 944–955. https://doi.org/10.1039/C6EM00222F

Kanakala, M., Pediredla, K., Pachiappan, N., Ramayi, R. R., & Rehman, T. (2019). Effect of Vitamin A supplementation on preventing recurrent acute lower respiratory tract infections in children. *International Journal of Contemporary Pediatrics*, *6*(4), 1632–1637. https://doi.org/10.18203/2349-3291.ijcp20192767

Khalil, I. A., Troeger, C., Rao, P. C., Blacker, B. F., Brown, A., Brewer, T. G., Colombara, D. V., De Hostos, E. L., Engmann, C., Guerrant, R. L., Haque, R., Houpt, E. R., Kang, G., Korpe, P. S., Kotloff, K. L., Lima, A. A. M., Petri, W. A., Platts-Mills, J. A., Shoultz, D. A., … Mokdad, A. H. (2018). Morbidity, mortality, and long-term consequences associated with diarrhoea from Cryptosporidium infection in children younger than 5 years: A meta-analyses study. *The Lancet Global Health*, *6*(7), e758–e768. https://doi.org/10.1016/S2214-109X(18)30283-3

Khan, J. R., Hossain, Md. B., Chakraborty, P. A., & Mistry, S. K. (2022). Household drinking water E. coli contamination and its associated risk with childhood diarrhea in Bangladesh. *Environmental Science and Pollution Research*, *29*(21), 32180–32189. https://doi.org/10.1007/s11356-021-18460-9

Khan, S. M., Bain, R. E. S., Lunze, K., Unalan, T., Beshanski-Pedersen, B., Slaymaker, T., Johnston, R., & Hancioglu, A. (2017). Optimizing household survey methods to monitor the Sustainable Development Goals targets 6.1 and 6.2 on drinking water, sanitation and hygiene: A mixed-methods field-test in Belize. *PLOS ONE*, *12*(12), e0189089. https://doi.org/10.1371/journal.pone.0189089

Kile, M. L., Rodrigues, E. G., Mazumdar, M., Dobson, C. B., Diao, N., Golam, M., Quamruzzaman, Q., Rahman, M., & Christiani, D. C. (2014). A prospective cohort study of the association between drinking water arsenic exposure and self-reported maternal health symptoms during pregnancy in Bangladesh. *Environmental Health, 13*(1), 13-29. https://doi.org/10.1186/1476-069X-13-29

Koolwal, G., and D. Van de Walle. 2013. "Access to Water, Women's Work, and Child Outcomes." *Economic Development and Cultural Chang,e* 61 (2): 369–405

Luby, S. P., Halder, A. K., Huda, T. M., Unicomb, L., Islam, M. S., Arnold, B. F., & Johnston, R. B. (2015). Microbiological contamination of drinking water associated with subsequent child diarrhea. *The American Journal of Tropical Medicine and Hygiene, 93*(5), 904–911. https://doi.org/10.4269/ajtmh.15-0274

Luby, S. P., Rahman, M., Arnold, B. F., Unicomb, L., Ashraf, S., Winch, P. J., Stewart, C. P., Begum, F., Hussain, F., Benjamin-Chung, J., Leontsini, E., Naser, A. M., Parvez, S. M., Hubbard, A. E., Lin, A., Nizame, F. A., Jannat, K., Ercumen, A., Ram, P. K., … Colford, J. M. (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: A cluster randomised controlled trial. *The Lancet Global Health, 6*(3), e302–e315. https://doi.org/10.1016/S2214-109X(17)30490-4

Modern, G., Mpolya, E., & Sauli, E. (2022). Causal relationship between environmental enteric dysfunction (EED), poor WaSH practices and growth failure in children from Rukwa-Tanzania. *Scientific African, 16,* e01281. https://doi.org/10.1016/j.sciaf.2022.e01281

Mulatu, T., Yimer, N. B., Alemnew, B., Linger, M., & Liben, M. L. (2021). Exclusive breastfeeding lowers the odds of childhood diarrhea and other medical conditions: Evidence from the 2016 Ethiopian demographic and health survey. *Italian Journal of Pediatrics*, *47*(1), 166. https://doi.org/10.1186/s13052-021-01115-3

National Nutrition Survey 2018. Key Findings Report. (n.d.). Retrieved December 27, 2022, from https://www.unicef.org/pakistan/reports/national-nutrition-survey-2018-key-findings-report

Null, C., Stewart, C. P., Pickering, A. J., Dentz, H. N., Arnold, B. F., Arnold, C. D., Benjamin-Chung, J., Clasen, T., Dewey, K. G., Fernald, L. C. H., Hubbard, A. E., Kariger, P., Lin, A., Luby, S. P., Mertens, A., Njenga, S. M., Nyambane, G., Ram, P. K., & Colford, J. M. (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: A cluster-randomised controlled trial. *The Lancet Global Health, 6*(3), e316–e329. https://doi.org/10.1016/S2214-109X(18)30005-6

Osarogiagbon, W. O., & Isara, A. R. (2018). Knowledge of acute respiratory infection in under-fives and homebased practices by their caregivers in an urban community in southern Nigeria. *African Journal of Thoracic and Critical Care Medicine, 24*(4). https://doi.org/10.7196/AJTCCM.2018.v24i4.188

Paulson, K. R., Kamath, A. M., Alam, T., Bienhoff, K., Abady, G. G., Abbas, J., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abd-Elsalam, S. M., Abdoli, A., Abedi, A., Abolhassani, H., Abreu, L. G., Abu-Gharbieh, E., Abu-Rmeileh, N. M., Abushouk, A. I., Adamu, A. L., Adebayo, O. M., … Kassebaum, N. J. (2021). Global, regional, and national progress towards Sustainable Development Goal 3.2 for neonatal and child health: All-cause and cause-specific mortality findings from the Global burden of disease study 2019. *The Lancet, 398*(10303), 870–905. https://doi.org/10.1016/S0140-6736(21)01207-1

Prüss-Ustün, A., Wolf, J., Bartram, J., Clasen, T., Cumming, O., Freeman, M. C., Gordon, B., Hunter, P. R., Medlicott, K., & Johnston, R. (2019). Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: An updated analysis with a focus on low- and middle-income countries. International *Journal of Hygiene and Environmental Health, 222*(5), 765–777. https://doi.org/10.1016/j.ijheh.2019.05.004

Rahman, M. M., Kunwar, S. B., & Bohara, A. K. (2021). The interconnection between water quality level and health status: An analysis of Escherichia Coli contamination and drinking water from Nepal. *Water Resources and Economics, 34*, 100179. https://doi.org/10.1016/j.wre.2021.100179

Rudan, I., O'Brien, K. L., Nair, H., Liu, L., Theodoratou, E., Qazi, S., Lukšić, I., Fischer Walker, C. L., Black, R. E., Campbell, H., & Child Health Epidemiology Reference Group (CHERG). (2013). Epidemiology and etiology of childhood pneumonia in 2010: Estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. *Journal of Global Health, 3*(1), 010401. https://doi.org/10.7189/jogh.03.010401

Santika, N. K. A., Efendi, F., Rachmawati, P. D., Has, E. M. M., Kusnanto, K., & Astutik, E. (2020). Determinants of diarrhea among children under two years old in Indonesia.

Children and Youth Services Review, 111, 104838. https://doi.org/10.1016/j.childyouth.2020.104838

Shrestha, S. K., Vicendese, D., & Erbas, B. (2020). Water, sanitation and hygiene practices associated with improved height-for-age, weight-for-height and weight-for-age z-scores among under-five children in Nepal. *BMC Pediatrics, 20*(1), 134. https://doi.org/10.1186/s12887-020-2010-9

Swarthout, J., Ram, P. K., Arnold, C. D., Dentz, H. N., Arnold, B. F., Kalungu, S., Lin, A., Njenga, S. M., Stewart, C. P., Colford, J. M., Null, C., & Pickering, A. J. (2020). Effects of individual and combined water, sanitation, handwashing, and nutritional interventions on child respiratory infections in rural Kenya: A cluster-randomized controlled trial. *The American Journal of Tropical Medicine and Hygiene, 102*(6), 1286–1295. https://doi.org/10.4269/ajtmh.19-0779

UNICEF (2018). Methodological Work—UNICEF MICS. Retrieved December 9, 2022, from https://mics.unicef.org/methodological-work

UNICEF (2022). Devastating floods in Pakistan. Retrieved November 26, 2022, from https://www.unicef.org/emergencies/devastating-floods-pakistan-2022

Usman, M. A., Gerber, N., & von Braun, J. (2019). The impact of drinking water quality and sanitation on child health: Evidence from rural Ethiopia. *Journal of Development Studies, 55*(10), 2193–2211. https://doi.org/10.1080/00220388.2018.1493193

VanDerslice, J., Popkin, B., & Briscoe, J. (1994). Drinking-water quality, sanitation, and breast-feeding: Their interactive effects on infant health. *Bulletin of the World Health Organization, 72*(4), 589–601. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2486614/

Wan, Zhengming, Hook, Simon, & Hulley, Glynn. (2021). MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V061 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD11A1.061

Wang, D., & Shen, Y. (2022). Sanitation and work time: Evidence from the toilet revolution in rural China. *World Development, 158,* 105992. https://doi.org/10.1016/j.worlddev.2022.105992

Walker, C. L. F., Perin, J., Katz, J., Tielsch, J. M., & Black, R. E. (2013). Diarrhea as a risk factor for acute lower respiratory tract infections among young children in low-income settings. *Journal of global health, 3*(1), 010402. https://doi.org/10.7189/jogh.03.010402

WorldPop, & Bondarenko, Maksym. (2020). Individual countries 1km population density (2000-2020) [Data set]. University of Southampton. https://doi.org/10.5258/SOTON/WP00674

WHO (2022). Action—Nutrition International—Pakistan—Vitamin A supplementation | Global database on the Implementation of Nutrition Action (GINA). Retrieved December 25, 2022, from *https://extranet.who.int/nutrition/gina/en/node/26182*

WHO (2022). Guidelines for Drinking-Water Quality: Fourth Edition Incorporating the First and Second Addenda. Retrieved March 9, 2023 (https://www.who.int/publications-detail-redirect/9789240045064).

WHO (2013) Micronutrient deficiencies. Retrieved from https://www.who.int/nutirtion/topics/vad/en/.

WHO & UNICEF (2022) Data of Joint Monitoring Program. Retrieved December 5, 2022, from https://washdata.org/data.

Zhang, J. (2012). The impact of water quality on health: Evidence from the drinking water infrastructure program in rural China. *Journal of Health Economics, 31*(1), 122–134. https://doi.org/10.1016/j.jhealeco.2011.08.008

Table 1. Mean comparison tests for child health outcomes with and without E. coli in household drinking water.

	Household POC water quality (E. coli)			
Variables	Contaminated	Uncontaminated	Total	t-test
Diarrhoea	0.187	0.139	0.178	-0.048***
Weight-for-height	-0.424	-0.324	-0.403	0.101***
Weight-for-age	-1.325	-1.109	-1.281	0.217***
ARI	0.099	0.086	0.096	-0.012*
Observations	10,056	2,615	12,671	

Notes: Authors' calculations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Test for the mean comparison between households consuming water uncontaminated and contaminated with E. coli.

Table 2. OLS estimates of the effects of drinking water quality on child health.

Dependent variables	Diarrhoea	Weight-for-height	Weight-for-age	ARI
	(1)	(2)	(3)	(4)
Water quality (1=E.coli present)	0.008	-0.067*	-0.076**	0.004
	(0.010)	(0.040)	(0.036)	(0.007)
Log of population	Yes	Yes	Yes	Yes
Night-time light	Yes	Yes	Yes	Yes
Rainfall	Yes	Yes	Yes	Yes
Temperature	Yes	Yes	Yes	Yes
Slope	Yes	Yes	Yes	Yes
Child, mother, HH head, and HH controls	Yes	Yes	Yes	Yes
District x Rural FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Clustered standard errors (robust) at the MICS cluster level are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. n = 12,671

Table 3. The IV estimates the effects of drinking water quality on child health.

	Water quality (1=E.coli present)			
Regression stage	(1)	(2)	(3)	(4)
Panel A. First-stage regression				
Community-level source water contamination (except self)	0.252*** (0.0165)	0.248*** (0.0165)	0.246*** (0.0165)	0.241*** (0.0165)
Control variables				
Child, mother characteristics	Yes	Yes	Yes	Yes
Household Characteristics		Yes	Yes	Yes
District Characteristics			Yes	Yes
District x rural FE				Yes
Year FE				Yes
Weak identification test				
Kleibergen-Paap rk Wald *F*-statistics	212.089			

Dependent variables	Diarrhoea	Weight-for-height	Weight-for-age	ARI
	(1)	(2)	(3)	(4)
Panel B. Second-stage regression				
Water quality (1=E. coli present)	0.079*	-0.312*	-0.322**	0.069**
	(0.044)	(0.171)	(0.151)	(0.035)
Control variables				
Child, mother characteristics	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes
District Characteristics	Yes	Yes	Yes	Yes
District x Rural FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Clustered standard errors (robust) at MICS cluster level are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The endogenous variable household POC water quality is instrumented by community-level source water contamination. n = 12,671

Table 4. Differential effects of drinking water quality on child health by age group

Variables	Younger than 6 months of age (2SLS)				Aged 6 months and older (2SLS)			
	Diarrhoea	Weight-for-Height	Weight-for-Age	ARI	Diarrhoea	Weight-for-Height	Weight-for-Age	ARI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Water quality (1= E. coli present)	-0.038	-0.356	0.015	0.012	0.094*	-0.349**	-0.358**	0.0771**
	(0.126)	(0.506)	(0.435)	(0.089)	(0.046)	(0.178)	(0.158)	(0.036)
Observations	1,270	1,270	1,270	1,270	11,401	11,401	11,401	11,401
Districts controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child, mother, HH head, and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District x Rural FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistics	47.3				206.4			

Clustered standard errors (robust) at MICS cluster level are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5. Differential effects of drinking water quality on child health depending on Vitamin A supplementation

Variables	Vitamin A-Supplemented group (2SLS)				Vitamin A Non- Supplemented group (2SLS)			
	Diarrhoea	Weight-for-Height	Weight-for-Age	ARI	Diarrhoea	Weight-for-Height	Weight-for-Age	ARI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Water quality (1= E. coli present)	-0.006	-0.159	-0.420*	0.023	0.221***	-0.569**	-0.325	0.150***
	(0.059)	(0.212)	(0.213)	(0.049)	(0.072)	(0.269)	(0.234)	(0.052)
Observations	5,855	5,855	5,855	5,855	5,546	5,546	5,546	5,546
Districts controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child, mother, HH head, and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District x Rural FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistics		127.5					98.41	

Clustered standard errors at MICS cluster level are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix Table 1. Summary Statistics

Variables	Mean	Std. Dev.	Min	Max
HH water quality (1= E. coli present)	0.793	0.404	0	1
Diarrhoea	0.178	0.382	0	1
Weight-for-height z-score	-0.403	1.360	-5	5
Weight-for-age z-score	-1.2816	1.301	-5.99	4.94
ARI	0.096	0.295	0	1
HH Demographic and Socioeconomic				
Child gender (1 = male)	0.514	0.499	0	1
Child age (month)	30.06	17.07	0	60
Child prenatal checks (4 or more)	0.609	0.487	0	1
Children under 5 years	2.399	1.329	1	10
HH head gender (1 = male)	0.924	0.263	0	1
HH head education	0.500	0.500	0	1
Maternal education	0.395	0.489	0	1
Maternal age (1 = < 20,…,4 = > 35)	2.559	0.847	1	4
Maternal smoking status	0.020	0.141	0	1
Wealth index quantile (1 = poorest,..., 5 = richest)	2.657	1.365	1	5
Number of HH members	9.391	4.920	2	42
Cooking place (1 = outdoor)	0.281	0.449	0	1
Flush toilet	0.469	0.499	0	1
Animal ownership	0.539	0.498	0	1
Residential area (1 = rural)	0.781	0.412	0	1
Year	2018	0.773	2017	2021
District level controls				
Log of population	5.782	1.481	1.107	10.847
Malaria prevalence rate	0.007	0.007	0.0001	0.037
Night-time light	1.242	5.440	0.072	48.730
Rainfall	515.594	380.485	44.860	1556.3
Temperature	31.09	6.188	8.902	41.979
Slope	6.963	8.706	0.281	34.355

Note: Sample size = 12,671

Appendix Table 2. Results of reduced-form regression for IV estimation

	Dependent Variable			
	Diarrhoea	Weight-for-height	Weight-for-age	ARI
Variables	(1)	(2)	(3)	(4)
Community source water quality (except self)	0.0190*	-0.0748*	-0.0770**	0.016**
	(0.010)	(0.0418)	(0.0365)	(0.00857)
Log of population	Yes	Yes	Yes	Yes
Night-time light	Yes	Yes	Yes	Yes
Rainfall	Yes	Yes	Yes	Yes
Temperature	Yes	Yes	Yes	Yes
Slope	Yes	Yes	Yes	Yes
Child, mother, HH head, and HH controls	Yes	Yes	Yes	Yes
District x rural FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Clustered standard errors (robust) at the MICS cluster level are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. n = 12,671.

Appendix Table 3. Sensitivity analysis excluding covariates from the 2SLS estimation model for the impact of drinking water quality on child health.

Variables	Diarrhoea			Weight-for-height		
	(1)	(2)	(3)	(4)	(5)	(6)
Water quality (1 = E. coli present)	0.077*	0.083*	0.079*	-0.346**	-0.346**	0.312*
	(0.042)	(0.042)	(0.044)	(0.16)	(0.167)	(0.171)
	Weight-for-age			ARI		
	(7)	(8)	(9)	(10)	(11)	(12)
	-0.301**	-0.296**	-0.322**	0.059*	0.060*	0.069**
	(0.149)	(0.148)	(0.152)	(0.034)	(0.034)	(0.035)
District characteristics			Yes			Yes
Household controls		Yes	Yes		Yes	Yes
Child, mother, HH head controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District x rural FE	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistics	225.1	220.9	212.0	225.1	220.9	212.0

Notes: Clustered standard errors (robust) at the MICS cluster level are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. n = 12,671

Appendix Table 4. Sensitivity analysis with and without WASH infrastructure and hygiene behaviour-related variables for 2SLS estimates of the impact of drinking water quality on child health

	Without water quality determinant variables		With water quality determinant variables	
	Water quality (1 = E. coli present)			
Panel A. First-stage regression	(1)		(2)	
Community-level source water contamination	0.241*** (0.0165)		0.239*** (0.0164)	
Panel B. Second-stage regression	Diarrhoea		Weight-for-height	
	(1)	(2)	(3)	(4)
Water quality (1 = E. coli present)	0.079*	0.074*	-0.312*	-0.326*
	(0.044)	(0.044)	(0.171)	(0.171)
	Weight-for-age		ARI	
	(5)	(6)	(7)	(8)
	-0.322**	-0.337**	0.069*	0.067*
	(0.152)	(0.153)	(0.035)	(0.035)
Piped water		Yes		Yes
Motorized pump/borehole		Yes		Yes
Tube well		Yes		Yes
Water treatment		Yes		Yes
Handwashing		Yes		Yes
Flush toilet	Yes	Yes	Yes	Yes
Animal ownership	Yes	Yes	Yes	Yes
Other controls (all)	Yes	Yes	Yes	Yes

Clustered standard errors (robust) at the MICS cluster level are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Access to flush toilets and animal ownership are included as control variables. Columns 1,3,5 and 7 in panel B presents the results of 2SLS similar to Table. 3. The regressions results presented in Columns 2,4,6 & 8 controls for potential factors that can significantly affect microbial water quality. n= 12,671.

Appendix Table 5: Impact of drinking water quality (measured as log of E. coli [CFU/100 ml]) on child health.

Variables	Diarrhoea (1)	Weight-for-height (2)	Weight-for-age (3)	ARI (4)
Water quality				
(Log [E. coli measured in CFU/100ml])	0.0147*	-0.058*	-0.059**	0.013**
	(0.008)	(0.031)	(0.028)	(0.006)
F-Statistics		362.3		
Controls (all)	Yes	Yes	Yes	Yes

Clustered standard errors (robust) at the MICS cluster level are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. n = 12,671.

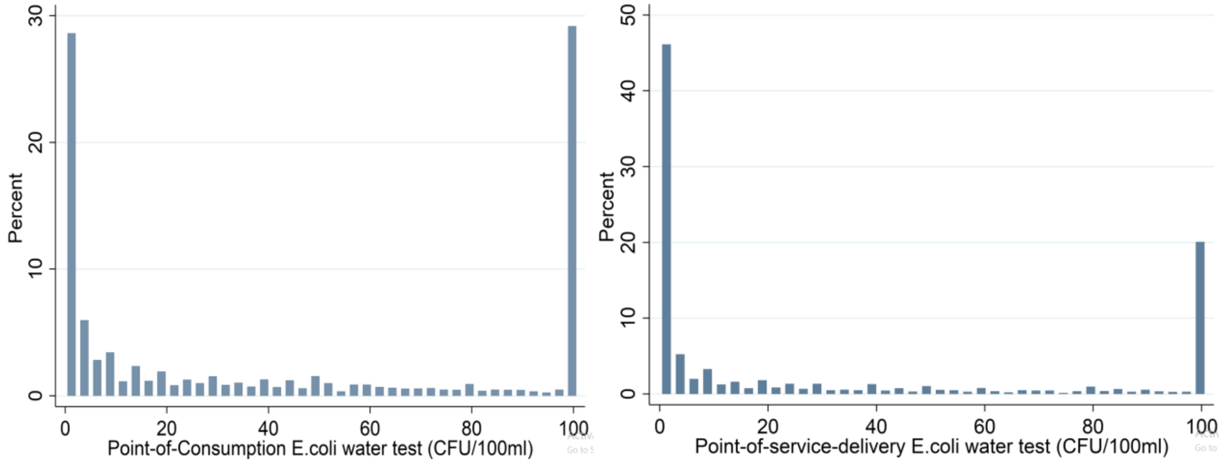

Figure 1. Point of consumption and source water quality status (E. coli [CFU/100 ml])

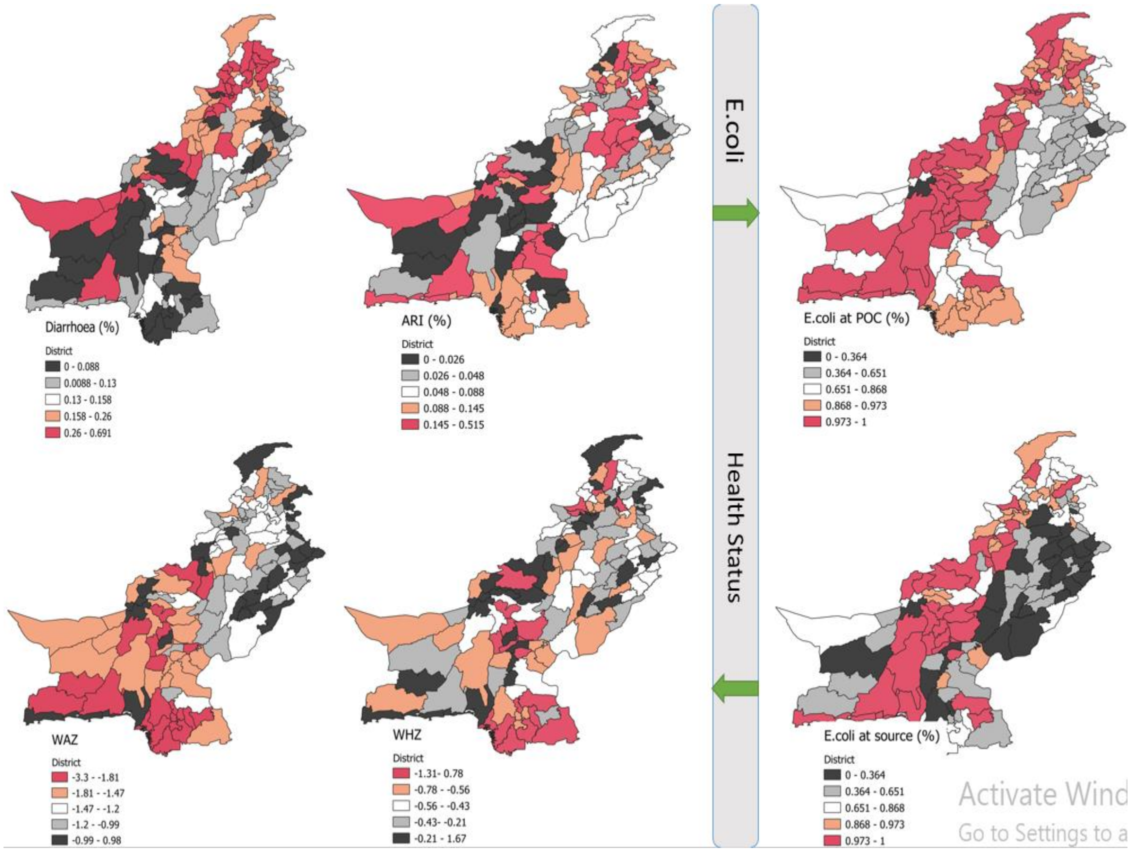

Figure 2. Microbial contamination and child health status at the district level

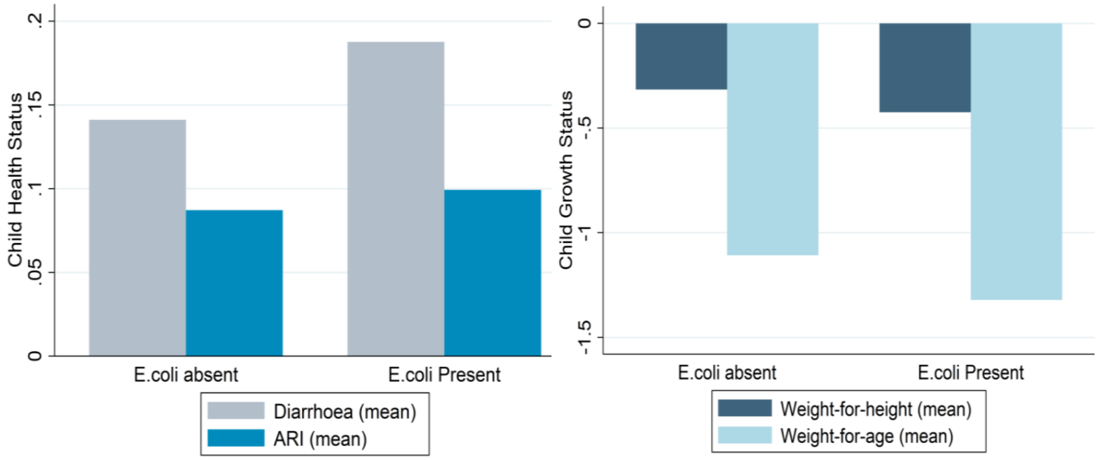

Figure 3. Relationship between child health indicators and E. coli contamination