Applied Spatial Analysis and Policy
https://doi.org/10.1007/s12061-023-09512-3

# Analyzing Effect of WASH Practices and District-Level Spatial Effects on Acute Respiratory Infections and Diarrhoea Among Under-Five Children in India

**Richa Vatsa[1] · Mukesh Ranjan[2] · Pravat Bhandari[3] · Ezra Gayawan[4]**

Received: 22 November 2021 / Accepted: 8 March 2023


**Abstract**
The United Nations sustainable development goal 6 (UN SDG-6) emphasizes equitable access to safe and affordable drinking water, sanitation, and hygiene (WASH) for all by 2030. Lack of WASH is health hazardous and hinders children's physical and educational development due to the frequent exposure to childhood illnesses. Previous researches from Africa have found strong linkages between WASH practices and childhood chronic undernutrition and related morbidities like diarrhoea, and acute respiratory infection (ARI). Furthermore, according to the National Health Profile of India, Report 2019, mortality from diseases among children is 27.2% for ARI and 10.5% for acute diarrhoea. The current study utilizes a Bayesian geoadditive modeling framework to explain the district-level spatial heterogeneity in cases of diarrhoea and ARI in India and simultaneously analyzes the effects of WASH practices and other socioeconomic covariates. The study results suggests that most districts situated in India's north and central regions had higher chances of ARI and diarrhoea and, cases of diarrhoea may reduce with the improved toilet facilities. Nevertheless, female children are less prone to ARI and Diarrhoea, whereas, stunted and wasted children are more susceptible to Diarrhoea only; young women with low education level are more likely to have children down with both the diseases. On the other hand, Hindu and ST have less while SC children have more chances of being sick with ARI and diarrhoea. Finally, the study may suggest to have an effective intervention of the Government for the identified regions of country with a high burden of ARI and Diarrhoea and a need for strategies for behavioural change in the people towards health and hygiene.

**Keywords** WASH indicators · ARI · Diarrhoea · Bayesian Geoadditive Regression · India

✉ Richa Vatsa
vatsa.richa@gmail.com

Extended author information available on the last page of the article

## Introduction

The United Nations sustainable development goal 6 (UN SDG-6) emphasizes equitable access to safe and affordable drinking water, sanitation, and hygiene (WASH) for all by 2030. WASH is a human right issue earlier recognized by the UN in 2010. It has a linkage with and ability to uplift the standard of living and ensures healthy lifestyle of any population. At the same time, lack of WASH is health hazardous and hinders children's physical and educational development because of the frequent exposure to childhood illnesses. WASH practices are seen as the most cost-effective strategy that reduces the burden of infant morbidity and mortality by limiting the transmission of pathogens through food, water, and environmental contamination (Velleman et al., 2014; Mascarini-Serra, 2011; WHO, 2011). However, according to WHO/UNICEF (2019), about 2.2 billion people worldwide do not have access to safe drinking water; 4.2 billion people do not have access to proper sanitation; 3 billion people do not use soap for handwashing, and almost 673 million people still practice open defecation.

In India, where nearly two-third of the population still lives in villages, most houses in rural places are closely situated and in clusters; localities are usually poorly managed with open drainage, no adequate sanitation and toilet facilities, unhygienic children stool disposal, and no proper facilities for safe drinking water (RGI & Census Commisioner, 2011; Chaudhuri et al., 2020; Borooah, 2022; Patil et al., 2014; Gopal et al., 2009; Pathak et al., 2015). According to NFHS-4 (2015-16) study, roughly 17% of households in India lacked a protected water facility, 61% utilized inappropriate sanitation, and 66% used unsafe kid stool disposal. The situation is further supplemented by lower literacy among women, lack of proper nutrition, and lack of good hygienic practices in handling children (IIPS & ICF, 2017).

Childhood chronic undernutrition and related morbidities like diarrhoea, fever, and acute respiratory infection (ARI) are major health problems in developing countries, especially South Asia, including India (Black et al., 2010). Although viral and bacterial infections are the primary cause of ARI, a lack of hygiene knowledge and increased pollution can aggravate and sometimes trigger the disease, particularly in children (Dey, 2019). The linkages between WASH practices with diarrhoea and ARI were reported from African countries (Luby & Halder, 2008; Kamm et al., 2014). Many Indian studies also found the role of hygiene and sanitation in triggering diarrhoeal and ARI diseases (Savitha & Gopalakrishnan, 2018; Selvaraj et al., 2014; Lakshminarayanan & Jayalakshmy, 2015; Singh & Singh, 2014). A study based on Jakarta (Krupnick et al., 1996) showed that the defence mechanism of mothers (washing hands after using the toilet) significantly reduced the chance of her child or herself getting diarrhoea. According to figures, 20% of total under-five death worldwide is due to ARI (Selvaraj et al., 2014); 13% of the overall disease burden of ARI was attributable to inadequate hygiene practices leading to 370,000 deaths in 2016 (Prüss-Ustün et al., 2019). On the other hand, more than 1400 under-five children die daily from diarrhoea, and out of that, 700 die every day due to a lack of WASH practices (UNICEF, 2021). As per the National Health Profile, 2019 report released by the Ministry of Health India showed that ARI contributes to most communicable diseases (nearly 70%), followed by acute diarrhoeal disease at 21.8% (MoHFW, 2019). Furthermore, mortality from diseases among children is 27.2% for ARI and 10.5% for acute diarrhoea (National Health Profile of India, Report 2019).

Recent findings based on factsheet data using aggregate level information of NFHS-5 suggest that diseases like diarrhoea in children under-five are more prevalent in rural areas (Prakash et al., 2021; Paul, 2020). Geographical variations in these childhood diseases have also been found in different regions, states, and districts of India, as it is in other developing countries (IIPS & ICF, 2017; Waghmare, 2020; Prakash et al., 2021; Dhirar et al., 2018; Gayawan et al., 2019). Given the threats caused by diarrhoea and respiratory illness worldwide, in-depth analyses of these two diseases become very important (Chakrabarti, 2012). Few studies (Goldman et al., 2002; Duraisamy, 2001; Krupnick et al., 1996) drawing data from developing economics have looked into the incidence of these two diseases and examined their determinants. In recent years, the modeling of childhood diseases has advanced shape because of the adoption of models that include covariates beyond socioeconomic factors and consider nonlinear, time-varying covariates, spatial patterns, and unobserved factors in a single framework.

Bayesian geoadditive model is one of the recently developed techniques which help draw meaningful conclusions of spatial patterns of childhood diseases. The model contains the well-established framework of generalized linear models and generalized additive models as special cases. It allows for the measuring of the effects of geographical location or spatio-temporal effects at the same platform. Studies, mainly from Africa, utilized individual-level information to understand the inequality in childhood diseases using advanced semiparametric geoadditive models, e.g., Khatab and Fahrmeir (2008) and Khatab et al. (2016) in Egypt; Kandala et al. (2009) in the Demographic Republic of Congo; Khatab (2014) in Egypt and Nigeria, and Gayawan et al. (2019) in several West African countries. Few studies from India have utilised aggregate-level data to study the spatial variations and contextual determinants of these childhood diseases; however, they completely ignored the individual-level information (Kamath et al., 2018; Ghosh et al., 2021). However, Yadav (2017) attempted to study childhood diseases using geoadditive modeling with individual-level NFHS-3 data but the spatial unit was limited to the state level, which can conceal variations at districts levels particularly within states.

An analysis of childhood illnesses across geographical locations needs to consider that events at proximate locations can be highly correlated especially in cultural settings where cultural practices and behaviour are similar across neighbouring locations. Models that ignore this important assumption and considered district effects as independent, as used in most previous studies in India, are capable of underestimating standard errors for the district factors (Goldstein, 2011; Rabe-Hesketh & Everitt, 2003). Indian states are bigger than few countries in their geographical dimension; therefore, it makes it very challenging to plan and implement policies at state level, more so that the development in India varies a lot at the district level. Hence, the district is a critical geographical unit to study childhood morbidity and mortality risk factors. The present study uses a Bayesian geoadditive modeling framework to explain the district-level spatial heterogeneity in childhood diseases like diarrhoea and ARI in India and simultaneously analyzes the effects of WASH practices and other socioeconomic covariates. The aim is to ascertain if there are clusters in the incidence of these diseases among the districts of India, to quantify the patterns of the effects of WASH practices on the diseases and to discern the functional relationship between the childhood diseases and continuous covariates including maternal and child age.

## Materials and Methods

### Data Source

Data used in the analysis were obtained from the most recent round of National Family Health Survey (NFHS)-4 (2015-16) which is available at www.DHSprogram.com. The NFHS-4 survey considered a stratified two-stage sampling plan. At the first stage, the primary sampling units (PSUs) were villages and Census Enumeration Blocks (CEBs) in rural and urban areas, respectively. The PSUs were selected according to the 2011 census sampling frame with probability proportional to size (PPS) sampling. At the second stage of sampling, the selected PSUs with an estimated population of at least 300 households were divided into groups of 100–150 households with the systematic sampling scheme.

The NFHS-4 collects data on population, health, and nutrition in each Indian state and union territory. Most importantly, it provides district-level estimates for a wide variety of important indicators. NFHS-4 gathered data from 601,509 households, including 699,686 women and 103,525 men. The data files are as follows: household, male, reproductive age female (or individual), couple, births, and kids. We utilized the kids-file (KR) for our work. In kids file, information on children born to a mother during the five years prior to the survey was retrospectively asked. Out of 259,627 under-five children, 223,507 from 636 Indian districts were considered for analysis, selecting data for the alive children, mother's BMI measured, and removing data for non-mainland union territories (Andaman Nicobar Islands and Lakshadweep). In the survey, the information about the symptoms of ARI and occurrence of diarrhoea was collected for two weeks before survey and it is given for all alive children born 2010 or later. Mothers reported that nearly 3% of children in the sample had symptoms of ARI. The symptoms of ARI in the data consists of cough accompanied by short rapid breathing that is chest related and / or difficult breathing that is chest related. Similarly, nearly 9% of the children in the sample had diarrhoea in the two weeks before survey as reported by their mothers. The entire analysis was carried out on 223,507 children.

### Variables

We considered the following independent and dependent variables. Dependent Variables: Symptoms of ARI two weeks before the survey and diarrhoea two weeks prior to the survey. ARI and diarrhoea were each re-categorised as dichotomous variable (0 "No"

/1 "Yes"). Independent variables with categorical form were: WASH indicators-Child stool disposal (safe/unsafe), toilet facility in household (improved / unimproved, according to WHO definition), and water facility (protected/unprotected, according to WHO definition). Other covariates included in model were sex of children (female/male), birth order into four categories (1/2/3/4 & above), stunting (low height for age) (yes/no), wasting (low weight for height) (yes/no), mother's BMI (underweight/overweight/ normal/ obese), Vitamin A intake (yes/no), mother's education (No education/Primary/Secondary/ Higher), wealth index (poorest/poorer/middle/richer/richest), type of residence (urban/ rural), caste (SC/ST/OBC/Others), and religion (Hindu/Muslim/Others). These independent variables were included in the model as dummy variables. Other predictor variables were, district (spatial covariate), mother's age and child's age (continuous covariates).

## Methods

We considered a geoadditive regression model for binary responses to model our dichotomous dependent variables given the above-mentioned categorical, continuous, and spatial covariates under study. The description of the model is as follows. Let $y_i$ denote the disease (ARI or diarrhoea) of an $i^{th}$ child, with dichotomous categories as 'yes' and 'no,' and $x_i$ be the vector of covariates that $y_i$ may be regressed on via a logistic model defined as

$$P\left[y_i = 1\right] = \frac{\exp\left(\eta_i\right)}{1+\exp\left(\eta_i\right)},\ i = 1 : n.$$

The term $\eta_i$ in a generalized linear model is restricted to linear predictor such that, $\eta_i = x_i{}'\beta$, with regression parameters β. However, linear or non-linear effects of covariates and variations due to spatial covariates in the response variable $y_i$ are very natural in practice. A Bayesian structured additive regression model (Umlauf et al., 2015) extends the strictly linear predictor to accommodate linear effects of categorical covariates, non-linear effects of continuous covariates, spatial effects of geographical divisions at the same platform with the predictor

$$\eta_i = u_i'\beta + f_1\left(z_i\right) + f_2\left(w_i\right) + g\left(S_i\right),$$

where, $u_i$ is a vector of categorical covariates for the $i^{th}$ child. The terms $z_i$ and $w_i$ denote the child's and mother's ages, respectively, and $S_i$ refers to the district as the spatial unit. The term $\beta$ represents the vector of linear effects. Notations $f_1(.), f_2(.)$ denote non-linear smooth effects for ages of the child and the mother, respectively. Further, the function $g(.)$ refers to the nonlinear spatial effects of districts.

The Bayesian approach was considered for inference and prior distributions were considered as follows. Non-informative prior was assumed for $\beta$, i.e., $P(\beta) \propto 1$. Unknown non-linear effects $f_r(.), r = 1 : 2$ were assigned Bayesian p-spline priors (Lang & Brezger, 2004). A penalized polynomial spline of degree $l$ was used to approximate $f_r\left(z_i\right)$ as a linear combination of $J = (m + l)$ B-spline basis functions, $B_j(.)$, i.e., $f_r\left(z_i\right) = \sum_{j=1}^{J} \gamma_j B_j(z_i), r = 1 : 2$, with the space of $z_i$ defined by the equally-spaced knots $z_{i,min} = \zeta_{i,0} < \zeta_{i,1} < \cdots < \zeta_{i,m} = z_{i,max}$. The precision of spline can be ensured by assuming a large number of knots. In Bayesian analogue to this case, the unknown parameters $\gamma_j$ were assigned with second-order random walk priors, i.e.,

$$\gamma_j = 2\gamma_{j-1} - \gamma_{j-2} + e_k, \text{with } e_j \sim N(0, \sigma_e^2).$$

The variance parameters $\sigma_e^2$ were assigned inverse-gamma prior with pre-specified hyper-parameters $a$ and $b$.

Furthermore, the unknown spatial effects g was modelled a priori with a Gaussian Markov random field (GMRF) (Rue & Held, 2005). GMRF priors assume that neighbouring spatial locations are correlated as they may exhibit similar effects. Given spatial effects at locations other than k of an $i^{th}$ child, $g(S_{i,-k})$, the conditional distribution of a latent spatial effect $g(S_{i,k})$ is defined as,

$$g(S_{i,k})|g(S_{i,-k}) \sim N\left(\frac{1}{|N(k)|}\sum_{l\epsilon N(k)}^{g} g(S_{i,l}), \frac{1}{|N(k)|\delta}\right), \delta \sim gamma(c, d),$$

with mean parameter equal to the average of neighbouring regions and the variance parameter inversely proportional to $|N(k)|$, where $N(k)$ is the vector of neighbouring regions of a spatial location $S_{i,k}$. A gamma prior with pre-specified hyper-parameters $c$ and $d$ over the smoothness parameter $\delta$ was assumed. Kneib and Fahrmeir (2006) may be assessed for more details on above-mentioned geoadditive regression modeling for categorical response variables.

### Data Analysis

Before modelling the data, we considered a multicollinearity check between the covariates through the variance inflation factor (VIF). The values of VIF for all covariates were found less than 2.5 indicating no multicollinearity between the considered covariates. Further, Markov chain Monte Carlo (MCMC) sampling techniques were utilized with the software package BayesX and R-package BayesXSrc (Brezger et al., 2005) for the Bayesian inference. To ensure the convergence of MCMC samples, we considered 50,000 MCMC iterations, with 5000 burn-in period, 10 steps, with default hyperparameter setting $a = b = c = d = 0.001$, and default degree of spline 3 with 20 equidistant knots in the spline. We also performed sensitivity analysis to investigate the impact of the hyperpriors on results by changing the values of the hyperparameters as (a,b) = (c,d) = (1, 0.0005) and (0.0005, 0.0005). We, further, used the deviance information criterion (DIC) (Spiegelhalter et al., 2002) to compare the fitted models with the different choices of hyper-parameters. In all, the estimates turned out to be indistinguishable, implying that the estimates are less sensitive to the choice of hyper-parameters. We, therefore, present results for the default values.

## Results

Table 1 presents the frequency analysis of background characteristics of the under-five children under study. Nearly 3% and 9% of children under age five years had symptoms of ARI and diarrhoea, respectively. In more than 65% of the households, unsafe practices of child stool disposal were found; about 17% of household had unprotected water facilities, while 61% used unimproved toilet facilities.

Nearly 38% of the under-five children were stunted, whereas 20.42% were wasted, and more than 50% had no vitamin A intake. More than 35% of the children were the first child, while 16% were the third or above child of their respective mothers. Around 52% of the children were male. The majority belonged to the OBC caste category and were mostly Hindu. More than 75% of the children lived in rural areas, and the majority belonged to the poorer and the poorest economic class. While the majority (~45%) of the mothers had completed their secondary education, nearly 31% were illiterate. More than 60% of the mothers had normal weight; nearly 24% were underweight (using the Asian classification of weight category).

## Linear Effects

Tables 2 and 3 present the posterior odds ratio and 95% credible intervals for the linear effects of categorical covariates on ARI and diarrhoea. Regarding toilet facilities, children from households with improved facilities had significantly lower odds of having diarrhoea compared with those from households with unimproved facilities. The

**Table 1** Frequency distribution of considered characteristics of children, their mothers and households

| Variables | | N (%) | Variables | | N (%) |
|---|---|---|---|---|---|
| ARI | Yes | 6404 (2.87) | Caste | SC | 42,393 (18.97) |
| | No | 217,103 (97.13) | | ST | 43,858 (19.62) |
| Diarrhoea | Yes | 20,680 (9.25) | | OBC | 88,502 (39.60) |
| | No | 202,827 (90.75) | | Others | 48,754 (21.81) |
| Stool Disposal | Safe | 75,468 (33.77) | Religion | Hindu | 162,314 (72.62) |
| | Unsafe | 148,039 (66.23) | | Muslim | 34,848 (15.59) |
| Water Facility | Protected | 185,730 (83.10) | | Others | 26,345 (11.79) |
| | Unprotected | 37,777 (16.90) | Wealth Index | Poorest | 58,125 (26.00) |
| Toilet Facility | Improved | 87,080 (38.96) | | Poorer | 52,939 (23.69) |
| | Unimproved | 136,427 (61.04) | | Middle | 44,824 (20.05) |
| Stunting | Yes | 85,830 (38.40) | | Richer | 37,365 (16.72) |
| | No | 137,677 (61.60) | | Richest | 30,254 (13.54) |
| Wasting | Yes | 45,647 (20.42) | Mother's Highest level of Education | No Education | 68,554 (30.73) |
| | No | 177,860 (79.58) | | Primary | 32,604 (14.61) |
| Vitamin A intake | Yes | 104,593 (46.80) | | Secondary | 101,161 (45.34) |
| | No | 118,914 (53.20) | | Higher | 20,806 (9.32) |
| Sex of Child | Female | 107,926 (48.29) | Mother's BMI | Underweight | 53,240 (23.82) |
| | Male | 115,581 (51.71) | | Normal | 139,190 (62.27) |
| Birth Order | 1 | 82,370 (36.86) | | Overweight | 26,684 (11.94) |
| | 2 | 69,314 (31.01) | | Obese | 4393 (1.97) |
| | 3 | 36,034 (16.12) | | | |
| | 4 and above | 35,789 (16.01) | | | |
| Place of Residence | Urban | 52,987 (23.71) | | | |
| | Rural | 170,520 (76.29) | | | |

estimates for other WASH indicators were not significant in both cases of ARI and diarrhoea. Compared with the male category, the female children had lower odds of getting down with ARI or diarrhoea. Compared with birth order four and above, the children of first birth order had a significantly less chance of falling sick with diarrhoea. Wasting and stunting in under-five children were significant causes to catch diarrhoea, while vitamin A intake had no significant role in preventing both diseases.

**Table 2** Posterior odds ratio and 95% credible intervals (CrI) for the linear effects of categorical covariates on ARI (left) and diarrhoea (right)

| Variable | | ARI | | Diarrhoea | |
|---|---|---|---|---|---|
| | | Odds ratio | 95% CrI | Odds ratio | 95% CrI |
| Toilet | Unimproved | 1 | | 1 | |
| | Improved | 1.012 | (0.98, 1.045) | 0.962* | (0.943, 0.981) |
| Water | Unprotected | 1 | | 1 | |
| | Protected | 0.971 | (0.938, 1.003) | 1.008 | (0.986, 1.03) |
| Stool Disposal | Unsafe | 1 | | 1 | |
| | Safe | 0.993 | (0.958, 1.029) | 0.984 | (0.966, 1.003) |
| Sex of Child | Male | 1 | | 1 | |
| | Female | 0.924* | (0.9, 0.947) | 0.961* | (0.947, 0.975) |
| Birth Order | 4 and Above | 1 | | 1 | |
| | 1 | 0.96 | (0.911, 1.01) | 0.955* | (0.927, 0.985) |
| | 2 | 0.969 | (0.926, 1.015) | 0.996 | (0.97, 1.023) |
| | 3 | 0.995 | (0.944, 1.049) | 1.018 | (0.987, 1.05) |
| Stunted | No | 1 | | 1 | |
| | Yes | 1.009 | (0.981, 1.037) | 1.022* | (1.005, 1.039) |
| Wasted | No | 1 | | 1 | |
| | Yes | 1.007 | (0.974, 1.041) | 1.043* | (1.025, 1.062) |
| Vitamin A intake | No | 1 | | 1 | |
| | Yes | 1.026 | (1.000, 1.055) | 1.015 | (0.999, 1.032) |
| Religion | Others | 1 | | 1 | |
| | Hindu | 0.924* | (0.875, 0.978) | 0.913* | (0.883, 0.944) |
| | Muslim | 1.067 | (0.995, 1.139) | 1.058* | (1.015, 1.101) |
| Place of Residence | Rural | 1 | | 1 | |
| | Urban | 0.987 | (0.95, 1.024) | 1.017 | (0.995, 1.038) |
| Caste | Others | 1 | | 1 | |
| | SC | 1.026 | (0.971, 1.079) | 1.044* | (1.012, 1.076) |
| | ST | 0.965 | (0.896, 1.042) | 0.926* | (0.888, 0.966) |
| | OBC | 0.998 | (0.952, 1.045) | 1.019 | (0.993, 1.046) |
| Wealth Category | Poorest | 1 | | 1 | |
| | Poorer | 1.03 | (0.973, 1.088) | 1.019 | (0.986, 1.053) |
| | Middle | 1.053* | (1.001, 1.108) | 1.009 | (0.978, 1.039) |
| | Richer | 0.989 | (0.929, 1.053) | 1.007 | (0.971, 1.044) |
| | Richest | 0.917* | (0.842, 0.996) | 0.933* | (0.889, 0.98) |

**Table 3** Continued posterior odds ratio and 95% credible intervals (CrI) for the linear effects of categorical covariates on ARI (left) and diarrhoea (right)

| Variable | | ARI | | Diarrhoea | |
|---|---|---|---|---|---|
| | | Odds ratio | 95% CrI | Odds ratio | 95% CrI |
| Mother's Education | Illiterate | 1 | | 1 | |
| | Primary | 1.095* | (1.035, 1.16) | 1.071* | (1.036, 1.109) |
| | Secondary | 1.054* | (1.009, 1.1) | 1.045* | (1.019, 1.072) |
| | Higher | 1.026 | (0.948, 1.111) | 0.975 | (0.931, 1.022) |
| Mother's BMI | Normal | 1 | | 1 | |
| | Underweight | 1.011 | (0.946, 1.08) | 1.008 | (0.968, 1.049) |
| | Overweight | 1.000 | (0.929, 1.072) | 1.001 | (0.958, 1.047) |
| | Obese | 1.045 | (0.911, 1.194) | 1.015 | (0.93, 1.104) |

The results further show that Hindu had less chances of getting ARI and diarrhoea, while Muslim children were more likely to suffer from diarrhoea as compared to those of other religions. Similarly, the children belonging to the SC caste were more vulnerable to diarrhoea; however, those from the ST caste categories were significantly immune to diarrhoea but the results for the OBC caste were not significant for both diseases. Further, compared to the rural areas urban residency had no significant role in the spread of ARI and diarrhoeal disease among the children. The richest were protected against both diseases; whereas, children from the middle wealth category were more exposed to ARI disease.

In connection with mothers' education, the odds for diarrhoea and ARI are higher among children whose mothers attained primary and secondary levels of education when compared with those whose mothers had no education and the estimates age significant. However, the estimates for the weight categories of the mothers measured by their BMI are not significant.

## Non-Linear Effects

Figures 1 and 2 presents the estimates for the nonlinear effects of child's and mother's ages, respectively, showing the posterior means (solid lines) and 95% credible intervals (dashed lines) for each of ARI and diarrhoea. The estimates for child's age for both illnesses present patterns that depict smooth inverted curves. The likelihood of ARI increased with the age of child and peaks near 7 months, thereafter it slowly decreased with advancement in age. Similar results were found for diarrhoea revealing that the chances of diarrhoea also surged with the age of child and reached a peak at about 9 months, followed by a sharp decline. On the other hand, likelihood of ARI and diarrhoea in children decreased slowly as their mothers advanced in age.

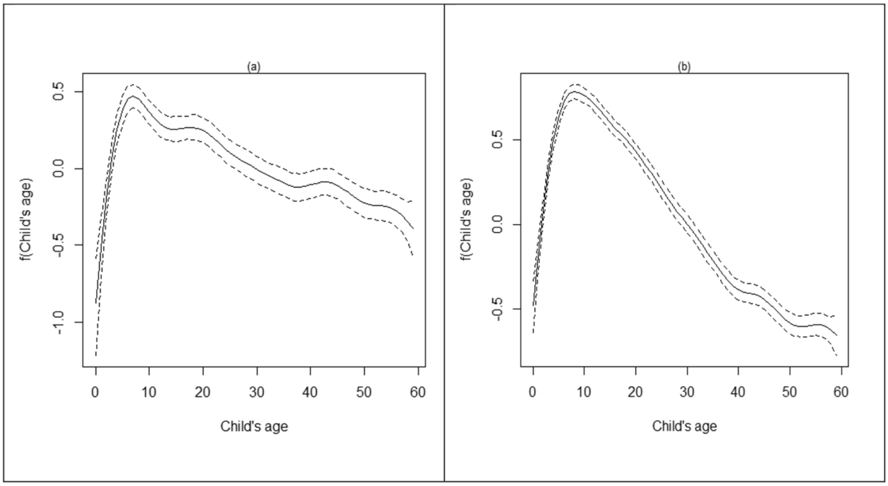


**Fig. 1** Non-linear effects of child's age in months on ARI (**a**) and diarrhoea (**b**)

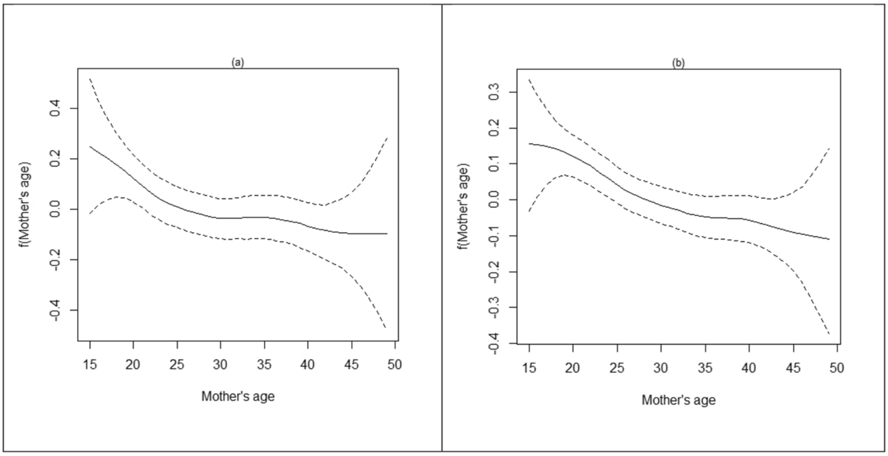


**Fig. 2** Non-linear effects of mother's age in years on ARI (**a**) and diarrhoea (**b**)

## Spatial Effects

Figure 3 presents the posterior means and 95% credible intervals for the spatial effects of the districts of India on under-five diseases ARI and diarrhoea. Posterior means are plotted on the left where red and green colours reflect positive and negative impacts, respectively. While 95% credible intervals of spatial impacts are shown in the right panel of the figure, with white and black colours indicating significant positive and negative impacts, respectively, grey colour represents districts with insignificant effects. Most districts of the states, Jammu and Kashmir,

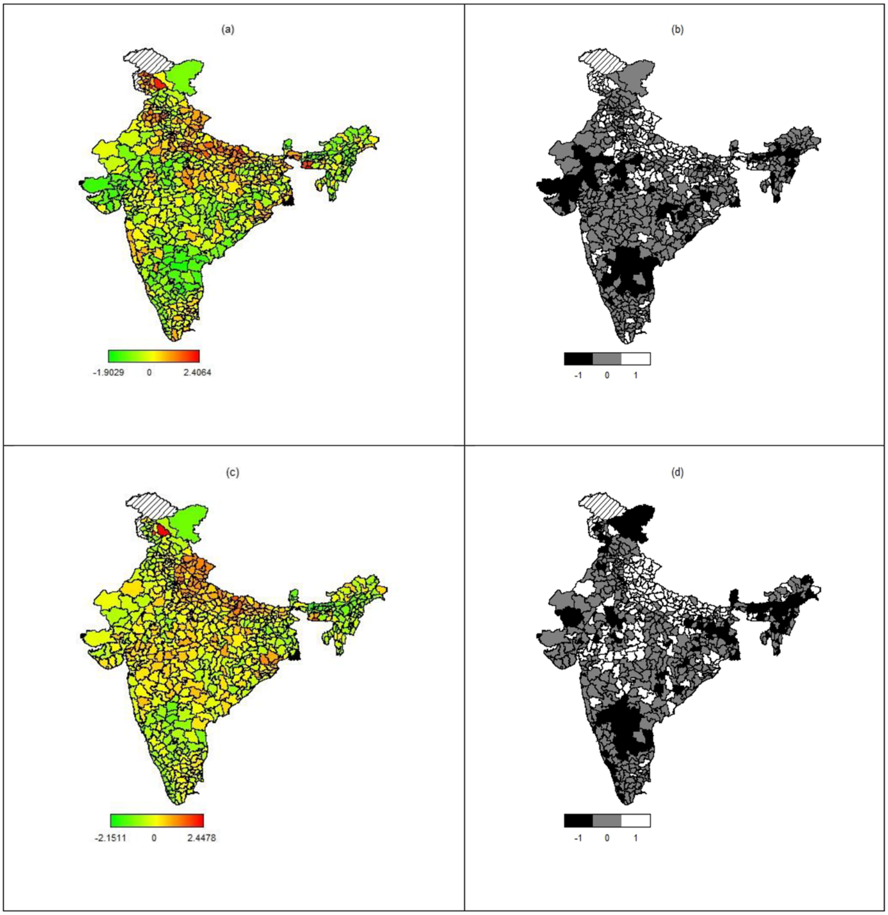


**Fig. 3** Mean spatial effects (left) and 95% credible interval of spatial effects (right) of different states of the country on ARI (**a**) and (**b**) and diarrhoea (**c**) and (**d**)

Punjab, Haryana, Delhi, Uttarakhand, Uttar Pradesh, and Meghalaya, had significant positive impact on ARI, implying higher likelihoods in these locations. In contrast, many districts of the states, Gujrat, Karnataka, Andhra Pradesh, Assam, Mizoram, Manipur and Nagaland had a significant adverse effect on ARI, pointing at low prospects of the disease in the regions.

On the other hand, many of the districts under the states of Kerala, Karnataka, Andhra Pradesh, West Bengal, Sikkim, Assam, Tripura, Manipur, and Nagaland had a significant negative effect on diarrhoea. Therefore, the children from these districts had less chances of developing diarrhoea in under-five children. Nevertheless, the most of districts with significant positive impacts on diarrhoea lie in the states, Uttarakhand, Uttar Pradesh, Bihar, Maharashtra, Madhya Pradesh, Telangana, Orissa, Tamil Nadu, and Meghalaya, indicating high chances of diarrhoeal cases in native children.

The majority of districts situated in the north and central regions of India showed high likelihoods of ARI and diarrhoea. At the same time, districts in southern, east, north-east, and western regions of India were found protected against diarrhoea and ARI.

## Discussion

Our study analyzed the effect of WASH practices and district-level spatial variations on the childhood diseases like ARI and Diarrhoea using a categorical Bayesian geo-additive modeling. The main benefit of the model is that it integrates spatial, linear, and non-linear effects at the same platform and gives realistic estimates of the coefficient of various covariates as it takes into account the possible association that may exist among neighbouring locations. All WASH indicators considered in the study had no significant effects on ARI and diarrhoea, except improved toilet facility that showed significant negative impact on diarrhoea. It is possible that the presence of other variables considered weaken the effects of these WASH indicators or that other factors over and beyond those covered in this study place some role. Similar to this finding, some studies that focussed on child health and unhealthy sanitary practices found higher odds of under-five mortality due to open defecation and unsafe stool disposal practices in India (Dwivedi et al., 2019). Toilet facilities also called for hygiene practices; generally, people with improved toilets wash their hands with soap, while people with no toilet or unimproved facilities usually refrain from good hand-hygiene practices. Irrespective of water quality, regular hand-washing with soap was found to halve the rate of diarrhoea (Lankester & Grills, 2019). It has been seen that families who start using soap and water often gradually stop again, so it would be necessary to adopt some behavioural change strategy to reinforce this in the longer run (Lankester & Grills, 2019; Gadgil et al., 2011; Ram, 2013). In our study sample, more than three-fourth of under-five children belong to rural areas, with majority being in the poorest or poorer wealth category, which may not consider it necessary to put in place some sanitary measures due to the perceived cost involved. Behavioural changes would therefore be necessary to get this set of people adopt hand washing using soap after toilet use.

Further, our study results indicate that wasting and stunting in under-five children positively affect diarrhoea only. However, most previous Indian studies found a significant association of malnutrition with both diarrhoea and ARI (Yadav et al., 2015). Our study highlights that only the richest had a significant negative effect on both diseases compared to the poorest wealth category. It is expected because the richest have more resources at their disposal. They try to take proper care of their children with a nutritious diet and hygienic practices. Further, as highlighted by several studies in different parts of developing countries (Gayawan et al., 2019; Fagbohungbe et al., 2020; Ghosh et al., 2021), educated women are able to utilize their knowledge to improve the care of their households than the less educated ones, leading to improved health condition for every household member. They would also be able to express themselves to health practitioners than the uneducated women when the need arises. However, the findings from our study show nonsignificant effects for higher level of education while the odds for both diseases are slightly higher among women who attained primary or secondary level of education. Such unexpected findings call for concern and there is the need for further investigation to unravel educational effects on child health among Indian women. The current study suggests that residing in a particular locality or region be it rural or urban area is not an independent factor affecting ARI and diarrhoeal diseases. However, there are studies that indicate higher prevalence of diarrhoea among under-five children living in rural areas (Ghosh et al., 2021). On the other hand, other socio-demographic and economic factors did play their roles in affecting the diseases. Findings based on the non-linear effects of children's age on diarrhoea and ARI are critical in the first ten months of life; after that, the positive effect of ages has a curvilinear decrement. Similar findings were reported by (Yadav, 2017) using NFHS-3 data at the state level. It might be because of numerous vaccinations scheduled for growing children, so that as they age, they become better at coping with the environment.

Similarly, the non-linear effects of mothers' ages 15 to 25 years had a significantly increasing impact on ARI and diarrhoea, but beyond age 25, it had a decreasing effect. This suggests that as the mothers aged, they gain more experience in disease prevention methods and thereby protecting their children. Ghosh et al. (2021) discovered similar results regarding the impact of mothers' age on diarrhoea in a study based on NFHS-4 data.

Furthermore, the current study suggests that most districts situated in India's north and central regions had higher chances of ARI and diarrhoea. In these regions, social and economic indicators have been much below the expectation, such as a lower literacy level among women, higher poverty, and sub-optimal performance of health facilities which could be individually or collectively some of the reasons for the findings. Ghosh et al. (2021) found that the hotspots for diarrhoea were mainly centred in the India's central and eastern coastal regions. A similar study (Chaurasia et al., 2020) found that diarrhoeal prevalence was highly concentrated in the northern region of India in certain parts of Madhya Pradesh, Chhattisgarh Jharkhand, and Bihar. Yadav (2017) also found a high prevalence of diarrhoea, fever, and ARI in the Eastern, Central, and northern states, while lower disease prevalence in southern states followed by western states.

Other community-based small studies in India have reported poor socio-economic factors, low level of literacy, malnutrition, and several other factors that contributed to increasing the burden of ARI among children (Acharya et al., 2003; Broor et al., 2001; Savitha et al., 2007). A study in Bangladesh for under-five children found that younger age of the child, male gender, undernutrition, lower maternal education, low wealth quintile, and poor breastfeeding practices were associated with diarrhoea (Leung et al., 2015; Yadav, 2017) also found that being underweight and belonging to the poorest quintile had a significantly higher association with diarrhoea, fever, and ARI.

## Limitations of the Study

The present study comes with certain limitations and therefore, interpretation of the study findings should be made with cautions. First, the information on childhood illness was based on self-reported accounts by the mothers, rather than diagnosed by a health investigator. Self-reported illness data may suffer from under- or over-reporting, as these are dependent on mother's recall and willingness to provide accurate information, which may lead to biased estimates. Although limiting the recall period to 2 weeks reduces some degree of such bias. Second, the definition of each childhood disease considered in the present study was broad; no distinction was made between the severity of these diseases (i.e., acute or persistent) due to non-availability of data. Third, we do not have data on mother's hygiene practices, such as washing hands before preparation of meals, during feeding and after toilet use, which could be potential risk factors associated with the disease transmission especially diarrhoea (Agustina et al., 2013; Sheth & Obrah, 2004). Another major limitation is that we did not account for climatic factors (e.g., rainfall, temperature) in our model, and effects of such important factors might have contributed to the residual confounding (Azage et al., 2017; Bhandari et al., 2020).

## Conclusions and Policy Implications

Using data from India's 2015-16 National Family Health Survey, the present study investigated spatial patterns of childhood morbidities of ARI and diarrhoea, adjusting for the effects of WASH indicators and other bio-demographic and socioeconomic variables either at individual or household level. Our analysis revealed that the distributional pattern of both the diseases is strongly spatially structured. Our findings further suggest that children of the districts located in central and northern India have a higher association with childhood infectious diseases such as ARI and diarrhoea compared with children living in the districts situated in other regions. These findings may be helpful for the policymakers in recognizing how the geographic impact of location influences the epidemiology of childhood diseases and may help them developing spatially targeted interventions for efficient resource allocation. The policymakers may also effectuate strategies for behavioural change of the people in the long run towards health and hygiene by intensifying the programs like *Swachh Bharat Abhiyan* (Clean India Mission) and Total Sanitation Campaign (TSC), particularly in the districts where risk of childhood diseases is greater. Furthermore, our results provided strong support for flexibly modelling of metrical covariates including child's and mother's age that clearly found to have non-linear effects on the outcomes. We find that the risk of infections increases sharply during the first year of postnatal life and thereafter it starts declining. Such age effect of child indicates the need to pay attention on the infant and young child feeding (IYCF) practices. In this regard, more emphasis should be placed on the WHO guidelines on exclusive breastfeeding for the first six months after childbirth, as well as IYCF practises, because it is well established that children with better nutrition are at a lower risk of infection.

**Authors' Contributions** RV and MR conceptualized the paper; RV conducted the statistical analysis; MR and RV wrote Introduction; RV wrote Methodology and Results; MR and PB wrote Discussion and Conclusion; EG critically evaluated the manuscript; RV prepared the final manuscript. All authors agreed on the final version of the manuscript.

**Data Availability** The study is based on a secondary data set of National Family Health Survey 2015-16. The data is available in the public domain and can be accessed from the official site of The DHS Program: Demographic and Health Surveys (https://dhsprogram.com).

## Declarations

**Competing Interests** The authors declare that they have no potential competing interest with respect to the research, authorship, and/or publication of this article.

## References


Acharya, D., Prasanna, K. S., Nair, S., & Rao, R. S. (2003). Acute respiratory infections in children: A community based longitudinal study in south India. *Indian Journal of Public Health, 47*(1), 7–13.

Agustina, R., Sari, T. P., Satroamidjojo, S., Bovee-Oudenhoven, I. M., Feskens, E. J., & Kok, F. J. (2013). Association of food-hygiene practices and diarrhea prevalence among indonesian young children from low socioeconomic urban areas. *BMC Public Health, 13*(1), 977.

Azage, M., Kumie, A., Worku, A., Bagtzoglou, C., & Anagnostou, E. (2017). Effect of climatic variability on childhood diarrhea and its high risk periods in northwestern parts of Ethiopia. *PLoS One, 12*(10), e0186933.

Brezger, A., Kneib, T., & Lang, S. (2005). BayesX: Analyzing Bayesian structural additive regression models. *Journal of Statistical Software*, *14*(11), 1–22. https://doi.org/10.18637/jss.v014.i11

Bhandari, D., Bi, P., Sherchand, J. B., Dhimal, M., & Hanson-Easey, S. (2020). Assessing the effect of climate factors on childhood diarrhoea burden in Kathmandu, Nepal. *International Journal of Hygiene and Environmental Health, 223*(1), 199–206.

Black, R. E., et al. (2010). Presentation for the child health epidemiology reference group of WHO and UNICEF. Global, regional, and national causes of child mortality in 2008: A systematic analysis. *The Lancet, 375*(9730), 1969–1987.

Borooah, V. K. (2022). Development, sanitation and personal hygiene in India. *The European Journal of Development Research*, *34*(1), 103–123.

Broor, S., Pandey, R. M., Ghosh, M., Maitreyi, R. S., Lodha, R., Singhal, T., & Kabra, S. K. (2001). Risk factors for severe acute lower respiratory tract infection in under-five children. *Indian Pediatrics, 38*(12), 1361–1369.

Chakrabarti, A. (2012). Determinants of child morbidity and factors governing utilization of child health care: evidence from rural India. *Applied Economics*, *44*(1), 27–37.

Chaurasia, H., Srivastava, S., & Singh, J. K. (2020). Does seasonal variation affect diarrhoea prevalence among children in India? An analysis based on spatial regression models. *Children and Youth Services Review, 118*, 105453.

Chaudhuri, S., Roy, M., & Jain, A. (2020). Appraisal of WaSH (Water-Sanitation-Hygiene) infrastructure using a Composite Index, spatial algorithms and Sociodemographic correlates in rural India. *Journal of Environmental Informatics, 35*(1), 1–22.

Dey, S. (2019). Acute respiratory infections top list of communicable diseases in 2018. https://timesofindia.indiatimes.com/india/acute-respiratory-infections-top-list-of-communicable-diseases-in-2018/articleshow/71844929.cms . Accessed 7 Sept 2021.

Dhirar, N., Dudeja, S., Khandekar, J., & Bachani, D. (2018). Childhood morbidity and mortality in India—Analysis of national family health survey 4 (NFHS-4) findings. *Indian Pediatrics, 55*(4), 335–338.

Duraisamy, Palanigounder. (2001). Health status and curative health care in rural India, No 78, NCAER Working Papers, National Council of Applied Economic Research. https://EconPapers.repec.org/RePEc:nca:ncaerw:78

Dwivedi, L. K., Banerjee, K., Jain, N., Ranjan, M., & Dixit, P. (2019). Child health and unhealthy sanitary practices in India: Evidence from recent round of national family health Survey-IV. *SSM-Population Health, 7*, 100313.

Fagbohungbe, T. H., Gayawan, E., & Orunmoluyi, O. S. (2020). Spatial prediction of childhood malnutrition across space in Nigeria based on pointreferenced data: An SPDE approach. *Journal of Public Health Policy*, *41*, 464–480.

Gadgil, M. A., Unicomb, L., Luby, S. P., & Ram, P. K. (2011, December). Consistent soap availability correlates with use of inexpensive soap products and improved handwashing in low-income households in Dhaka, Bangladesh. In *American Society of Tropical Medicine and Hygiene Annual Meeting* (pp. 4–8).

Gayawan, E., Aladeniyi, O. B., Oladuti, O. M., Olopha, P., & Adebayo, S. B. (2019). Investigating the spatial patterns of common childhood morbidity in six neighboring West African Countries. *Journal of Epidemiology and Global Health, 9*(4), 315–323.

Ghosh, K., Chakraborty, A. S., & Mog, M. (2021). Prevalence of diarrhoea among under five children in India and its contextual determinants: A geo-spatial analysis. *Clinical Epidemiology and Global Health, 12*, 100813.

Goldman, N., Pebley, A. R., & Gragnolati, M. (2002). Choices about treatment for ARI and diarrhea in rural Guatemala. *Social Science & Medicine*, *55*(10), 1693–1712.

Goldstein, H. (2011). *Multilevel statistical models* (922 vol.). Wiley.

Gopal, S., Sarkar, R., Banda, K., Govindarajan, J., Harijan, B. B., Jeyakumar, M. B., & Balraj, V. (2009). Study of water supply & sanitation practices in India using geographic information systems: Some design & other considerations in a village setting. *Indian Journal of Medical Research, 129*, 233–241.

International Institute for Population Sciences (IIPS) and ICF. (2017). *National Family Health Survey (NFHS-4), 2015-16: India*. IIPS.

Kamath, A., Shetty, K., Unnikrishnan, B., Kaushik, S., & Rai, S. N. (2018). Prevalence, patterns, and predictors of diarrhea: A spatial-temporal comprehensive evaluation in India. *BMC Public Health, 18*(1), 1288.

Kamm, K. B., et al. (2014). Associations between presence of handwashing stations and soap in the home and diarrhoea and respiratory illness, in children less than five years old in rural western Kenya. *Tropical Medicine & International Health, 19*(4), 398–406.

Kandala, N. B., Emina, J. B., Nzita, P. D. K., & Cappuccio, F. P. (2009). Diarrhoea, acute respiratory infection, and fever among children in the Democratic Republic of Congo. *Social Science & Medicine, 68*(9), 1728–1736.

Khatab, K. (2014). Bayesian geoadditive mixed latent variable models with applications to child health problems in Egypt and Nigeria. *Advanced techniques for modelling maternal and child health in Africa* (pp. 49–81). Springer.

Khatab, K., & Fahrmeir, L. (2008). Analysis of childhood morbidity with geoadditive probit and latent variable model: a case study for Egypt. *American Journal of Tropical Medicine and Hygiene, 81*(1):116–28. https://doi.org/10.5282/ubm/epub.2505

Khatab, K., Adegboye, O., & Mohammed, T. I. (2016). Social and demographic factors associated with morbidities in young children in Egypt: A Bayesian geo-additive semi-parametric multin-omial model. *PLoS One, 11*(7), e0159173.

Kneib, T., & Fahrmeir, L. (2006). Structured additive regression for categorical Space–Time Data: A mixed Model Approach. *Biometrics, 62*(1), 109–118. https://doi.org/10.1111/j.1541-0420.2005.00392.x

Krupnick, A., Alberini, A., McGranahan, G., & Eskeland, G. S. (1996). Determinants of diarrhoeal disease in Jakarta. *World Bank Policy Research Working Paper*, No.1568, 1–33.

Lakshminarayanan, S., & Jayalakshmy, R. (2015). Diarrheal diseases among children in India: Current scenario and future perspectives. *Journal of Natural Science Biology and Medicine, 6*(1), 24–28.

Lang, S., & Brezger, A. (2004). Bayesian p-splines. *Journal of Computational and Graphical Statistics, 13*, 183–212.

Lankester, T., & Grills, N. J. (2019). *Setting up Community Health Programmes in Low and Middle Income Settings* (4th edn). Oxford Medicine Online. https://play.google.com/books/reader?id=IMKKDwAAQBAJ&pg=GBS.PA279&hl=en . Accessed 7 Sept 2021.

Leung, D. T., et al. (2015). Concurrent pneumonia in children under 5 years of age presenting to a diarrheal hospital in Dhaka, Bangladesh. *The American Journal of Tropical Medicine and Hygiene, 93*(4), 831–835.

Luby, S. P., & Halder, A. K. (2008). Associations among handwashing indicators, wealth, and symptoms of childhood respiratory illness in urban Bangladesh. *Tropical Medicine & International Health, 13*(6), 835–844.

Mascarini-Serra, L. (2011). Prevention of soil-transmitted Helminth infection. *Journal of Global Infectious Diseases, 3(2),* 175–182. https://doi.org/10.4103/0974-777X.81696

Ministry of Health & Family Welfare (MoHFW) (2019). National Health Profile. http://www.cbhidghs.nic.in/showfile.php?lid=1147. Accessed 7 Sept 2021.

National Health Profile (2019). 14th issue. Government of India.

Patil, S. R., Arnold, B. F., Salvatore, A. L., Briceno, B., Ganguly, S., Colford Jr, J. M., & Gertler, P. J. (2014). The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: A cluster randomized controlled trial. *PLoS Medicine, 11*(8), e1001709.

Pathak, B., Sharma, A., & Chattopadhyay, S. (2015). Sanitation Practices and Policies in India: Exploring determinants and their interlinkages. *South Asia Journal of Experimental Biology, 5*(6), 192–204.

Paul, P. (2020). Socio-demographic and environmental factors associated with diarrhoeal disease among children under five in India. *BMC Public Health, 20*(1), 1886.

Prakash, S., Gunnal, G. S., & Singh, N. M. (2021). Diarrhoea in children under 5 more prevalent in rural India: NFHS-5. *DownToEarth*. https://www.downtoearth.org.in/news/health/diarrhoea-in-children-under-5-more-prevalent-in-rural-india-nfhs-5-75058. Accessed 7 Sept 2021.

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 Volume, 222*(5), 765–777. https://doi.org/10.1016/j.ijheh.2019.05.004

Rabe-Hesketh, S., & Everitt, B. (2003). *Handbook of statistical analyses using Stata*. Chapman and Hall/CRC.

Ram, P. (2013). Practical guidance for measuring handwashing behavior: 2013 update. *Washington, DC: Water and Sanitation Program*.

Registrar General of India (RGI) & Census Commissioner (2011). Census of India 2011, Provisional Population Totals Paper 2, volume 1. https://censusindia.gov.in/2011-prov-results/paper2/data_files/india/paper2_1.pdf. Accessed 7 Sept 2021.

Rue, H., & Held, L. (2005). *Gaussian Markov random fields: Theory and applications*. CRC Press.

Savitha, A. K., & Gopalakrishnan, S. (2018). Determinants of acute respiratory infections among under five children in a rural area of Tamil Nadu, India. *Journal of Family Medicine and Primary Care, 7*(6), 1268.

Savitha, M. R., Nandeeshwara, S. B., Kumar, M. P., & Raju, C. K. (2007). Modifiable risk factors for acute lower respiratory tract infections. *The Indian Journal of Pediatrics, 74*(5), 477–482.

Selvaraj, K., Chinnakali, P., Majumdar, A., & Krishnan, I. S. (2014). Acute respiratory infections among under-5 children in India: A situational analysis. *Journal of Natural Science Biology and Medicine, 5*(1), 15.

Sheth, M., & Obrah, M. (2004). Diarrhea prevention through food safety education. *The Indian Journal of Pediatrics, 71*(10), 879–882.

Singh, A., & Singh, M. N. (2014). Diarrhoea and acute respiratory infections among under-five children in slums: Evidence from India (No. e208v1). *PeerJ.* PrePrints 2:e208v1. https://doi.org/10.7287/peerj.preprints.208v1

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B, 64*(Part-4), 583–639.

Umlauf, N., Adler, D., Kneib, T., Lang, S., & Zeileis, A. (2015). Structured additive regression models: An R interface to BayesX. *Journal of Statistical Software, 63*(21), 1–46.

Velleman, Y., Mason, E., Graham, W., Benova, L., Chopra, M., Campbell, O. M., Gordon, B., Wijesekera, S., Hounton, S., Esteves Mills, J., Curtis, V., Afsana, K., Boisson, S., Magoma, M., Cairncross, S., & Cumming, O. (2014). From joint thinking to joint action: A call to action on improving water, sanitation, and hygiene for maternal and newborn health. *PLoS Medicine, 11*, e1001771. https://doi.org/10.1371/journal.pmed.1001771

Waghmare, A. (2020). Respiratory infections in children are up: National Family Health Survey. https://www.business-standard.com/article/current-affairs/respiratory-infections-among-children-up-in-last-four-years-nfhs-data-120121401484_1.html . Accessed 4 Sept 2021.

Yadav, A. (2017). Spatial Variation of Childhood Diseases in India: An application of Geoadditive Bayesian Semiparametric Model. In *2017 International Population Conference*. IUSSP.

Yadav, A., Ladusingh, L., & Gayawan, E. (2015). Does a geographical context explain regional variation in child malnutrition in India? *Journal of Public Health, 23*(5), 277–287.

WHO. (2011). Guide to ship sanitation. *Controlling infectious diseases in the environment* (3rd edition, 8). World Health Organization. https://www.ncbi.nlm.nih.gov/books/NBK310829/. Accessed 4 Sept 2021.

WHO/UNICEF (2019). Progress on household drinking water, sanitation and hygiene 2000–2017: Special focus on inequalities.

WHO/UNICEF (2021). *The measurement and monitoring of water supply, sanitation and hygiene (WASH) affordability: a missing element of monitoring of Sustainable Development Goal (SDG) Targets 6.1 and 6.2*. UNICEF.



## Authors and Affiliations

**Richa Vatsa[1] · Mukesh Ranjan[2] · Pravat Bhandari[3] · Ezra Gayawan[4]**

[1] Department of Statistics, Central University of South Bihar, Gaya, India

[2] Department of Statistics, Mizoram University, Pachhunga University College Campus, Mizoram, Aizawl, India

[3] International Institute for Population Sciences, Mumbai, India

[4] Department of Statistics, Federal University of Technology, Akure, Nigeria