Int Health 2018; **10**: 324–332
doi:10.1093/inthealth/ihy037 Advance Access publication 30 May 2018

Groundwater use and diarrhoea in urban Nepal: novel application of a geostatistical interpolation technique linking environmental and epidemiologic survey data

Sadhana Shrestha[a], Takashi Nakamura[a], Jun Magome[a], Yoko Aihara[b], Naoki Kondo[c], Eiji Haramoto[a], Bikash Malla[a], Junko Shindo[a] and Kei Nishida[a,*]

[a]Interdisciplinary Centre for River Basin Environment, University of Yamanashi, 4-3-11 Takeda, 400-8511 Kofu, Japan; [b]Kobe Gakuin University, 518 Arise Ikawadani-Cho Nishi-ku, Kobe City, 651-2180, Japan; [c]School of Public Health, The University of Tokyo, 7-3-1 Hongo Bunkyo, 113-8654, Tokyo, Japan

*Corresponding author: Tel: +81-55-220-8593; E-mail: nishida@yamanashi.ac.jp

Received 19 February 2018; revised 6 April 2018; editorial decision 24 April 2018; accepted 24 April 2018

Background: Groundwater is a common domestic water source in developing countries, but is persistently contaminated with enteropathogens. However, studies on determinants of diarrhoea have predominantly focused on piped water. This study examines the relationship between groundwater microbial quality and household diarrhoea occurrence (HDO).

Methods: Considering it as a proxy of enteropathogens, this study analysed *Escherichia coli* concentrations in groundwater wells. Ordinary kriging, a geostatistical technique in geographic information systems, was used to interpolate the *E. coli* concentration to survey points that had secondary survey data (n=942). The relationship between *E. coli* and HDO using simple and multivariate statistical analyses in SPSS was analysed.

Results: A total of 77% of households used groundwater. One-third of households were without piped-water access (PWA), and these households were significantly more likely to use groundwater than those with PWA. Of the 87 households that reported HDO, 77% were groundwater users. Of the groundwater users, the households with HDO consumed groundwater with significantly higher *E. coli* concentrations than the households without HDO. Of the households without PWA, the increase in the *E. coli* concentration increased the odds of HDO (adjusted odds ratio=3.15; 95% CI=1.07–9.22).

Conclusion: It is suggested that the groundwater microbial quality is a risk factor for HDO and illustrates this by an application of an interpolation technique relevant for developing countries.

Keywords: Domestic water source, Faecal contamination, Groundwater users, Piped-water supply, Public health, Urban area

Introduction

Groundwater constitutes half of the potable water used in countries such as Bangladesh, China, India, Indonesia, Nepal, the Philippines, Thailand and Vietnam,[1] particularly for the poorest urban households.[2] While an important source of consumable water, groundwater has widely been reported to have faecal contamination in countries at all levels of economic development.[3–9] In the Kathmandu Valley, which is a representative area of a developing country in Nepal, groundwater is used by 52% of the households and is a major water source.[10] A total of 3% of groundwater users use it for drinking and 88% use it for bathing purposes.[11] Shrestha et al.[12] recorded numerous studies conducted at different times and showed that the faecal contamination of shallow groundwater is a persistent problem in the valley. This discussion provided evidence that shallow groundwater is unprotected and cannot be considered to be an improved source as defined by the World Health Organization (WHO). In addition, the presence of *Escherichia coli* along with other pathogens and viruses[13] points towards a serious threat of waterborne diseases to public health in the valley.

According to the WHO, nearly 1.7 billion cases of diarrhoeal disease occur each year and it is the leading cause of malnutrition and the death of approximately 525 000 children under the age of 5 years each year worldwide. It is the second leading cause of years of life lost due to premature death in Nepal[14] and the under-5 mortality rate is 35 per 1000 people.[15] To establish preventive measures for diarrhoea, countless studies have been performed to identify its determinants over the last few decades. Numerous studies have emphasized piped water as an important determinant—access,[16] use and quantity.[17] In an effort to reach Millennium Development Goal 7 Target10c, the coverage of improved drinking water sources increased to 90% in South Asia. The developments of improved drinking water sources (90% in South Asia) basically refer to other improved water sources (65% coverage), rather than piped water (25%).[18] Based on the data collected by the Asian Development Bank (ADB) in the Kathmandu Valley, 21% of households did not have access to a piped-water supply in 2009,[17] which rose to 34% in 2015.[11] In addition, the households that have access receive this supply for six or fewer hours per week.[11] Despite the fact that alternative sources represent a huge share in household water consumption, these sources have been analysed in few studies[19] as determinants of diarrhoea.

In developing countries, conducting large-scale surveys is difficult due to institutional and technical reasons, which leads to insufficient data for scientific studies. In this situation, the application of a geographic information system (GIS) to interpolate data spatially could be an affordable solution. Most of the studies on risk factors associated with diarrhoeal diseases are based on a questionnaire survey. Combining survey data with microbial concentration data can provide rigorous findings. Interdisciplinary studies combining environment and health are difficult to conduct in developing country settings due to technical limitations. If this study can link separate datasets of different disciplines, it will be a cost-effective and easy approach to obtain significant findings.

With the background of wide usage and the poor microbial quality of groundwater, as well as the lack of investigations regarding their relationship with diarrhoea, this study analyses the microbial quality of the groundwater as a determinant of diarrhoea in the Kathmandu Valley. The first part of this paper illustrates the spatial distribution of the groundwater microbial concentration using a geostatistical interpolation technique in GIS to integrate environmental data (microbial quality) with social data. An ordinary kriging interpolation is applied, which uses the statistical properties of the measured points to quantify the spatial autocorrelation between measured points and account for the spatial configuration of the sample points around the prediction location. The second part examines the potential impact of the groundwater microbial quality on household diarrhoea occurrence (HDO) with the aim to emphasize the necessity of pollution control interventions for water sources in addition to piped water.

Methods

Study area

The Kathmandu Valley lies in the central hilly region of Nepal and comprises 85% of the Kathmandu district, the entire Bhaktapur district and 50% of the Lalitpur district. The valley is situated at an elevation of 1300–1400 m above sea level, and is drained by the Bagmati River and its tributaries. The residents of the valley exploit two aquifers to meet their water needs—shallow and deep aquifers. The shallow aquifer is composed of up to 50 m of Quaternary arkosic sand, with some discontinuous interbedded silt and clay, and lies in the northernmost part of the valley.[20] The deep aquifer is beneath the interbedded clay and lignite aquitard of up to 200 m thickness and consisted of Pliocene sand and gravel, with interbedded lignite, peat and clay, and lies in the southern zone of the valley.[21] Based on the water table map and the piezometric surface elevation, the groundwater flow is towards the centre from the periphery.[22] The shallow aquifer is directly recharged via precipitation and small rivers, whereas the deep aquifer is recharged via precipitation from the northeast part of the valley.[23] Groundwater is tapped from the shallow aquifers via dug wells and tube wells, and from the deep aquifer via deep tube wells.

The Kathmandu Valley is the largest urban centre in the country; it has a population of 2.51 million people and shows an annual growth rate of 3.65%. The Kathmandu district alone has a population density of 4408 persons per km^2, which is 30 times higher than the national population density. Most of the wards (administrative units) of the Kathmandu metropolitan city (KMC) and the Lalitpur sub-metropolitan city (LSMC) have population densities higher than 10 000/km^2, especially the old settlements that lie in the centre, which have densities >20 000/km^2.[24] Kathmandu Upatyaka Khanepani Limited (KUKL) supplies 106 million litres/day (MLD) and 76 MLD of water during the wet and dry seasons, respectively, countering a water demand of 320 MLD, and the water loss due to leakage is 40%.[25] Of the alternative water sources used, 52% of the households use groundwater.

Groundwater sample collection and microbial analysis

In this study, 36 groundwater wells were randomly and uniformly selected from the wards in KMC and LSMC (Figure 1). The samples were collected once during August–September 2009, which was the wet season. The daily rainfall data of the sampling period was compared with those of August–September in 2008 and in 2010. The rainfall amount and pattern in the sampling period were similar to those in the previous and following years, indicating an absence of extreme rainfall events within the sampling window. Before collecting water samples, tube wells and dug wells that were fitted with a hand pump were purged for 1–2 min. For open dug wells, the water samples were brought to the surface using the same method used by the owners. The autoclaved 250 mL polythene bottles were rinsed with the groundwater to be sampled, then the samples were collected and immediately stored in an icebox.

E. coli was selected as the faecal indicator bacterium, and the concentrations in the groundwater previously reported[12,13] were used for statistical analysis. The EPA-approved IDEXX Quanti-Tray method (USA) using the Colilert reagent was applied. After mixing the reagent, 100 mL of the sample water was poured into a tray, and the tray was sealed and incubated for 24 h at 35±0.5°C. After incubation, the wells in the tray producing a blue colour under UV light were counted. The most probable number (MPN) table is referred to here to determine the MPN of the *E. coli* in 100 mL of the sample. Only one sample

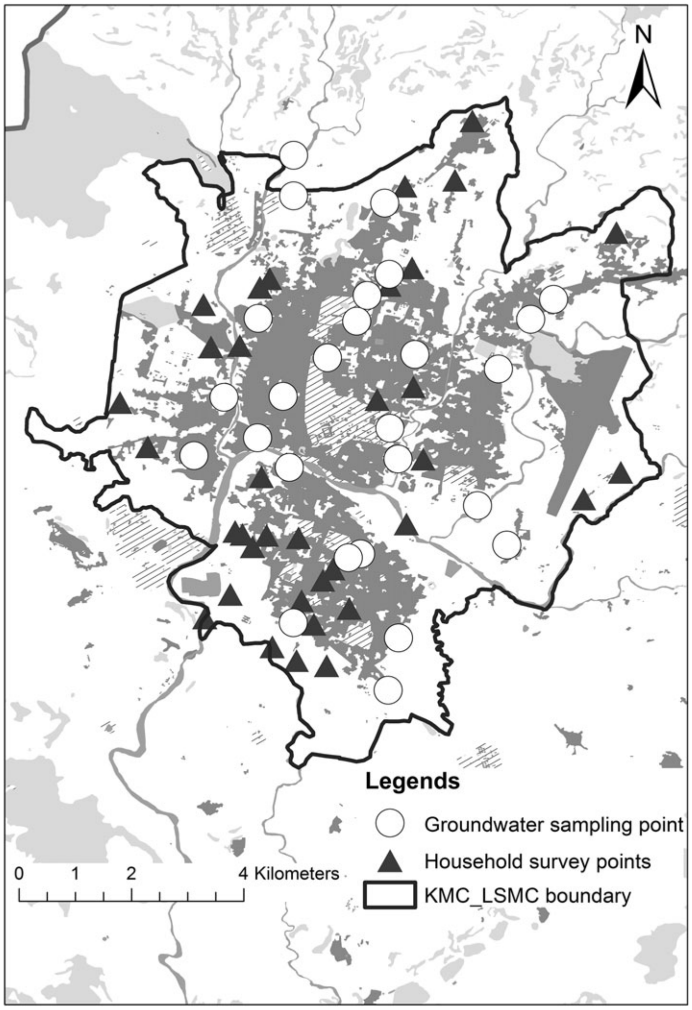

Figure 1. Locations of groundwater wells and questionnaire survey points in the Kathmandu metropolitan city (KMC) and the Lalitpur sub-metropolitan city (LSMC) area. Open circles (N_{well}=36) represent wells, and closed triangles (N_{survey_points}=35) represent survey points. The dark black boundary represents the boundary between KMC and LSMC. The river flowing from east to west separates KMC and LSMC.

was analysed per site, and field blanks, laboratory blanks and duplicates were not used.

In this study, *E. coli* was considered to be the faecal indicator bacterium to infer the presence of pathogens, and the Quanti-Tray method, which was used here, cannot discriminate between pathogenic and non-pathogenic strains of *E. coli*. *E. coli* is applicable for predicting total bacteria pathogen[26] and pathogenic viruses.[27] *E. coli* concentration has been widely applied in risk assessment studies in the form of faecal indicator ratio conversion.[28] Based on these lines of evidence, *E. coli* was treated as a proxy of pathogenic microorganisms to examine the impact of groundwater microbial quality on HDO.

Secondary data from household surveys

Secondary data were obtained from a household survey of the Kathmandu Valley Water Distribution, Sewerage and Urban Development Project conducted by the ADB from August to September in 2009.[10] The study used a multi-stage cluster survey design.[10] The ADB chose 84 geographical locations randomly and then drew a circle with each randomly generated geographical location at the centre. Furthermore, they selected 20 houses closest to the random location and all the households that were in those houses were interviewed. In the Kathmandu Valley, one house may contain more than one household. The total number of households interviewed was 2214. This study selected 35 geographical points among the 84 that lie within KMC and LMSC (Figure 1) for a total number of households of 942. Details of the survey strategy, and inclusion and exclusion criteria of households have been presented elsewhere.[10,17]

The secondary data included information on diarrhoea occurrences in at least one family member in the past month, which was defined as an HDO in this study. According to the WHO, diarrhoea is defined as the passage of three or more loose or liquid stools per day (or a more frequent passage than is normal for the individual); this definition was applied in the present study. In addition to health, information on water-use behaviour was also obtained from the secondary data for the statistical analysis. The data on the sources of water use and the purpose of use of each source, as well as socio-economic and demographic data, were also used.

Data analysis

Geostatistical spatial interpolation

Here, the ArcGIS Geostatistical Analyst Extension of the ArcMap software (Version 10.1, Environmental System Research Institute, USA) was used to interpolate the groundwater microbial quality data over KMC and LSMC, and to the 35 studied geographical points. Ordinary kriging was chosen as an appropriate tool based on previous groundwater studies.[29] Ordinary kriging weights surrounding the measured values were used to derive a prediction for each location, and the weights were based on the distance between the measured and predicted locations, as well as on the spatial arrangement of the measured points. Kriging requires spatial autocorrelation, where close samples are considered to be more similar than those farther apart. Based on this concept, an interpolated microbial concentration at each point was then assigned to all 20 houses associated with that point.

First, the spatial dependency was quantified using a semivariogram.[30] The empirical semivariogram is then fitted to a theoretical model, and the best-fit model of the spherical, exponential, linear and Gaussian models was used for the interpolation. The prediction accuracy of the fitted model was determined using the root mean square error (RMSE) and the coefficient of determination (R^2) obtained via cross-validation. The RMSE should be as small as possible.[29]

Statistical analysis

A χ^2 test of independence was used to compare the HDO between groundwater users and non-users. An independent t-test and its non-parametric alternative, the Mann–Whitney

U-test, were performed to compare the *E. coli* concentrations between HDO and without HDO in different groups:

- groundwater users;
- households that use groundwater only for drinking;
- households that use groundwater only for bathing;
- socio-demographic and water use characteristics;
- those with and without piped-water access (PWA).

The relationship between PWA and groundwater use was examined using the χ^2 test. Finally, to examine the relationship between the groundwater microbial quality and HDO, a simple logistic regression model was applied separately for the households with and without PWA. Each model was adjusted for common influencing socio-demographic characteristics, e.g. ethnicity, income, education of the decision maker, the head of the household and family size. Statistical significance was set at a p-value of <0.05. The statistical program IBM SPSS Statistics Version 21.0 (IBM Corporation, Armonk, NY, USA) was used for all statistical analyses.

Results

Observed groundwater *E. coli* concentration

E. coli was detected in 61% (22/36) of the groundwater samples.[12,13] The minimum, maximum and mean (±SD) values of the observed *E. coli* concentrations were <0, 3.84 and 1.271 (±1.274) Log$_{10}$ MPN/100 mL, respectively.

Interpolation of groundwater *E. coli* concentration and integration with survey data

Observed *E. coli* concentrations were used for the interpolation and, because the data best fit to a log-normal distribution, it was log transformed beforehand. The ordinary kriging method was used for interpolation. Different semivariogram models were tested for model prediction—circular, spherical, exponential, gaussian, k-Bessel and stable. Of the other semivariogram models, the gaussian model had the lowest RMSE (0.86) and the highest R^2 (0.41). Therefore, it was used to prepare a prediction map of the *E. coli* concentration (Figure 2) over KMC and LSMC. The interpolated *E. coli* concentration ranged from 0.00 Log$_{10}$ MPN/100 mL to 2.94 Log$_{10}$ MPN/100 mL. The predicted spatial distribution suggests that the *E. coli* concentrations gradually increase moving from northeast to southwest.

Using the prediction surface (Figure 2), the *E. coli* concentrations at 35 survey points (Figure 1) were determined. Of the 35 survey points, the mean (±SD), minimum and maximum *E. coli* concentrations were 1.707 (±0.796), 0.114 and 2.678 Log$_{10}$ MPN/100 mL, respectively.

Because the interviewed houses were the closest buildings to the survey points and because all the households inside a house were interviewed, all households that belong to one survey point were assigned the same interpolated *E. coli* concentration for the respective survey point.

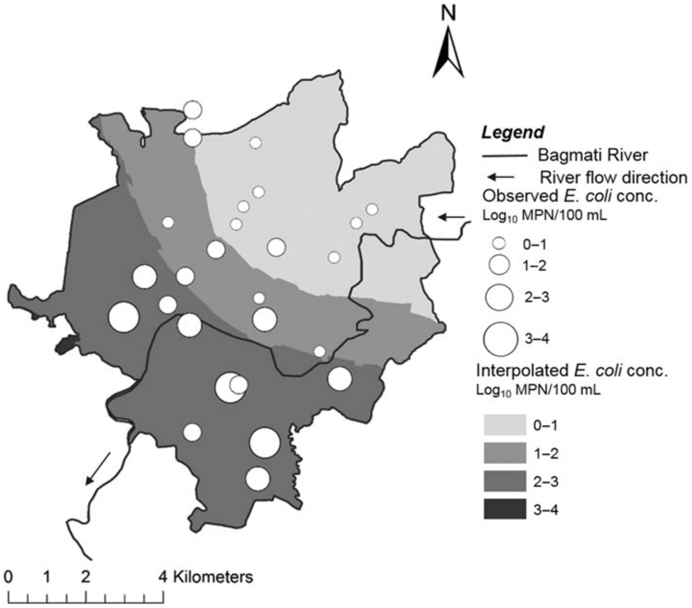

Figure 2. Observed and predicted surface of the *E. coli* concentration in the groundwater in the KMC and LSMC areas. The white bubbles represent the observed concentration, and the surface represents the interpolated *E. coli* concentration.

Descriptive information on socio-demographic and water using characteristics and HDO

Of the surveyed households, 47 (442/939), 52 (485/939) and 1 (12/939)% of households were from Brahmin/Chettri, Janajati and Dalit ethnicity, respectively, and 48% (366/762) of households had monthly incomes of >147 US$, 32% (242/762) had monthly incomes of 49–147 US$ and 20% (154/762) had monthly incomes of <49 US$ (US$ 1=NRs 101.94, as of 30 December 2017). A total of 73% (676/927) of the household heads were educated to above secondary level, 10% (97/927) to below secondary level and 17% (154/927) were illiterate. The average family size was 4.5 (±1.8).

A total of 75% (709/942) had PWA within their premises, 73% (684/942) households used groundwater and 29% (271/942) bought jar or bottled water. Of the groundwater users, 32% of households (184/575) bought jar water. Of the 942 households, 52% and 31% used piped water only for drinking and bathing, respectively. Likewise, 6% and 30% used groundwater only for drinking and bathing, respectively. Households without access to PWA had significantly higher (p-value <0.05) proportions of groundwater users compared with households that had PWA (Table 1). Likewise, the proportions of households that used groundwater only for drinking or for bathing were significantly greater in households without PWA than in those with PWA (Table 1). However, a significant difference between these two groups regarding other water sources such as vendor tankers or stone spouts was not observed.

A total of 9% (87/942) of the total households, 10% (68/684) of the groundwater users and 8% (21/258) of the groundwater nonusers reported HDO. Even though there was no statistical significance, the tendency to HDO seemed to be higher in groundwater users compared with non-users. Of the households reporting HDO, 77% (67/87) used groundwater.

Table 1. Proportion of groundwater users among households with and without PWA

Water-use characteristics		Piped-water access**		p-value
		No	Yes	
Groundwater use	No	23^a	234^c	<0.05#
	Yes	209^b	475^d	
Drinking only groundwater	No	173^a	707^c	<0.05*
	Yes	59^b	2^d	
Bathing only in groundwater	No	98^a	556^c	<0.05#
	Yes	134^b	153^d	

Proportion of groundwater users in: without PWA=b/(a+b); with PWA=d/(c+d).
#, χ^2 test; *, Fisher's exact test; **, out of 942 households, one household has missing information on PWA.

Relationship between HDO and groundwater microbial quality

Groundwater users

Groundwater users were selected for further analysis and the *E. coli* concentration was compared between users with HDO and without HDO across different household characteristics (Table 2). Table 2 shows an apparently higher mean *E. coli* concentration in the HDO group compared with the without HDO group across all household characteristics except for Dalit ethnicity. However, a significant difference was achieved for groundwater user households, where the household head's education level was below secondary level or was illiterate, households with no PWA and households in which members bathed only in groundwater. However, the *E. coli* concentrations were not significantly different between the HDO and without HDO groups among the households that drank only groundwater and among those that bathed in all types of water.

Households without PWA

Among households without PWA, the *E. coli* concentration was significantly (p-value <0.05) higher in the HDO group than in the without HDO group. However, this relationship was not evident in households that had PWA. In a simple logistic regression analysis, the *E. coli* concentration in the groundwater was significantly and positively associated with HDO (Table 3). With a one unit (Log_{10} 1 MPN/100 mL) increase in groundwater *E. coli* concentration, the odds of HDO increased three-fold (adjusted odds ratio [AOR]=3.15; 95% CI=1.07–9.22; p-value=0.03). Another factor that was associated with HDO was education. Compared with households with well-educated household heads, those with less-educated or illiterate household heads were at 4.6-fold (AOR=4.59; p-value=0.05) and 5.3-fold (AOR=5.3; p-value=0.03) more risk of HDO, respectively. Conversely, neither the *E. coli* concentration nor the groundwater use was significantly associated with HDO in households with PWA. The factors that were associated with HDO were ethnicity and the education level of the household head.

Discussion

The aim of this study was to identify the association of groundwater microbial quality with HDO by applying interpolation tools to link the groundwater microbial data (environmental data) with epidemiologic survey data. A total of 61% of the observed groundwater samples in this study detected *E. coli* and exceeded the National Drinking Water Quality Guideline, which states that *E. coli* should not be present in collected samples. These results are in line with previous groundwater studies in the Kathmandu Valley.[13,31] These comparative findings indicated persisting faecal contamination in the aquifer and confirmed that the water sources in most of the parts of study area fell under the 'unimproved source' category as defined by the WHO. As discussed in Shrestha et al.,[12] the possible sources of the microbial contamination of the groundwater wells in this urban area could be anthropogenic, e.g. leaky sewers, improperly constructed septic tanks, and improperly designed and managed storm drains that carried a mixture of sewage and storm water, which frequently overflowed, polluting the land surface.

These experimental *E. coli* concentration data were used for a geostatistical interpolation applying GIS for integration with the survey data. Even though the observed high *E. coli* concentration values were underestimated and the low values were overestimated, which is an inherent feature of interpolation algorithms,[32] the *E. coli* interpolation revealed that the quality of water worsened from northeast to southwest within the study area. JICA[33] (1990) divided the Kathmandu Valley watershed into three groundwater districts—northern (NGD), central (CGD) and southern (SGD). The northeastern part of the study area lies in the NGD, and the rest lies in the CGD. Previous studies have suggested higher chemical contamination in CGD compared with in NGD.[33,34] Most of the groundwater wells in CGD had a high ammonia, high nitrogen content[33] and high electrical conductivity compared with those in NGD.[34] The sediments in NGD consist of highly permeable sand and gravel, and exhibit the greatest groundwater potential, whereas those in CGD are comprised of very thick impermeable black clay, and a shallow layer, up to a depth of 20 m, overlies the thick clay. Such differences in the permeability of the sediments could produce a better dilution effect of the precipitation on the chemical, as well as the microbial contaminants in NGD, given the similar subsurface pollution.

Groundwater is the most commonly used alternative water source:

- 75% (709/941) households had PWA;
- 73% (684/942) households used groundwater;
- 29% (271/941) bought jar water.

Possessing a groundwater well within a compound makes households use it more often (ADB 2010).[10] It was also observed in this survey data that, compared with households that had PWA, those without access were more likely to use groundwater, and more likely to rely solely on groundwater for drinking and bathing purposes (Table 1), with 10% of

Table 2. Comparison of the *E. coli* concentration between households with and without diarrhoea occurrence among different strata of groundwater users

Groundwater users (N=684)	Diarrhoea frequency		*E. coli* concentration in groundwater Log_{10} MPN/100 mL		p-value
	No	Yes	Diarrhoea no	Diarrhoea yes	
Groundwater users	617	67	1.58	1.79	0.05
Ethnicity					
Brahmi/Chettri	299	39	1.60	1.83	0.09
Janajati	309	26	1.56	1.71	0.35
Dalit	7	1	2.05	1.64	1.00*
Household income#					
>US$ 147	245	22	1.61	1.87	0.09
US$ 49–147	162	20	1.55	1.87	0.11
<US$ 49	94	10	1.53	1.81	0.33
Education of household head					
Above secondary level	451	38	1.60	1.66	0.65
Below secondary level	66	9	1.64	2.10	0.047*
Illiterate	91	18	1.48	1.86	0.042*
Water treatment					
No	132	10	1.48	1.81	0.20
Yes	479	56	1.61	1.77	0.16
Piped-water access					
No	168	10	1.58	2.07	0.008*
Yes	449	57	1.58	1.72	0.23
Drinking only groundwater					
No	558	65	1.61	1.78	0.10
Yes	59	2	1.37	1.90	0.45
Bathing only in groundwater					
No	350	47	1.60	1.72	0.35
Yes	267	20	1.57	1.95	0.035*

*Mann–Whitney U-test; #, US$ 1=NRs 101.94, as of 30 December 2017.

groundwater using and 8% of non-using households reporting HDO, respectively. Similarly, 77% (70/87) of the records of HDO were groundwater users. These higher tendencies of HDO among groundwater users give a weak indication of some connection between groundwater use and HDO.

Across all household characteristics among groundwater users, the *E. coli* concentration was apparently higher in the HDO group than in the without HDO group. The *E. coli* concentration was significantly higher in the HDO groups than in the without HDO groups among the groundwater users and among the households that use groundwater only for bathing. However, such a significant difference was not observed among those households that drank only groundwater, but bathed in all types of water. In this study, 25% of the households did not have PWA. Among these households, the *E. coli* concentration was significantly higher in households with HDO than those without HDO. Among the households without PWA, the likelihood of HDO increased by three-fold with a log unit increase in *E. coli* concentration (Table 3). However, such an association was not significant among households with PWA. The association of the *E. coli* concentration with HDO under different conditions indicates that the microbial contamination of groundwater could be an important factor aggravating diarrhoea in the study area; this association was prominent when PWA was lacking.

A review of studies on global outbreaks and epidemiology, and studies on pathogens and risk assessment demonstrated that consumers served by an untreated groundwater supply remain at risk for enteric diseases.[35] Previous risk assessment studies on enteric infections from a variety of pathogens in groundwater estimated high risk and advised proper risk management strategies.[36] The only risk assessment study in the Kathmandu Valley[31] estimated a high risk of diarrhoea from groundwater and suggested the implementation of risk reduction programmes. Comparable with this finding of a high *E. coli* concentration among the 'bathing only in groundwater' group, Shrestha et al.[31] estimated a high risk of diarrhoea via the bathing pathway. During swimming and bathing, people usually accidentally ingest water and children ingest higher amounts than adults.[37] Bathers in polluted seawater are exposed to an increased risk of gastrointestinal infections.[38] These outcomes have opened an avenue

Table 3. Factors associated with household diarrhoea occurrence (HDO)

Factors	Frequency (N=232)	B	p-value	Exp(B)	95% CI for Exp(B)	
					Lower	Upper
E. coli concentration in groundwater (Log_{10} MPN/100 mL)	1.68 (±0.80)	1.22	0.03	3.15	1.07	9.22
Groundwater use						
No	23	Ref.				
Yes	209	−0.51	0.56	0.60	0.11	3.33
Ethnicity						
Brahmin/Chettri	119	Ref.				
Janajati	109	−1.13	0.09	0.32	0.09	1.18
Dalit	4	−20.01	1.00	0.000	0.000	0.000
Household income*						
>US$ 147	82	Ref.				
US$ 49–147	68	−0.31	0.68	0.74	0.17	3.19
<US$ 49	44	−0.07	0.94	0.94	0.19	4.64
Education of household head						
Above secondary level	156	Ref.				
Below secondary level	34	1.52	0.05	4.59	1.01	20.90
Illiterate	39	1.67	0.03	5.29	1.16	24.08
Family size (people)	4.5 (±1.8)	−0.06	0.73	0.94	0.66	1.33

B, beta coefficient; Exp(B), adjusted odds ratio; CI, confidence interval; *, US$ 1=NRs 101.94, as of 30 December 2017; Ref., reference group.

for researching the role of bathing in disease transmission, especially among young children. Overall, with a huge proportion of residents relying heavily on shallow groundwater, the findings of this study strongly indicate a potential risk from contaminated groundwater to public health, as well as suggesting the immediate application of groundwater pollution control approaches.

The biggest limitation of the present study is the lack of microbial data for other water sources. However, the observational data were compared with published data of piped, stored[39] and stone spout water.[40] The mean (558) and maximum (6867) values of *E. coli* (MPN/100 mL) in groundwater are higher than those in piped (mean=44; maximum=488), stone spout (mean=107; maximum=500) and stored (mean=301; maximum=2420) water. This comparison indicates that groundwater is a highly polluted major water source. However, it is strongly recommended that the microbial qualities of other water sources, together with groundwater, should be analysed to obtain a more realistic picture in future studies. The second limitation is the use of the faecal indicator bacterium, *E. coli*; the use of enteropathogens is recommended for quantitative microbial analyses, even though it was not feasible in this study due to limited resources and time. Data on pathogens such as protozoa are being accumulated in continuing research,[13] and further analyses will be reported in the future. The third limitation is the moderate regression accuracy (R^2=0.41) of the interpolation method, which could be due to the small sample size. Small sample size is thought to be a factor that leads to reduced accuracy. Therefore, using a larger sample size in future studies is recommended. Moreover, the recall period of HDO was 1 month in this study; this may have led to a recall bias. A previous study showed a 42–44% drop in diarrhoea reporting when the recall period was only 1 week.[40] In this study, it is highly likely that the respondents might have forgotten the frequency of diarrhoea, which might have resulted in an underestimation of HDO. In addition, single water sampling instead of daily or weekly sampling to characterize water quality in the wet season is another limitation of this study. A previous study[12] showed consistency in the *E. coli* detections and concentrations in wet seasons of different years and showed significant variations between dry and wet seasons in a single year. Despite these limitations, this study is an example of an integration of environmental with social data, which is especially important for developing countries that face financial restraints to conducting wider scale and interdisciplinary studies.

Conclusions

This study successfully integrated environmental data with survey data to perform an interdisciplinary analysis. With the application of the geostatistical interpolation technique, the observed *E. coli* concentration of groundwater samples in the wet season of 2009 was estimated for 35 survey points that had survey data of the same season and year. The *E. coli* concentration was significantly higher in the HDO group compared with the without HDO group among groundwater users and among households that use only groundwater for bathing. A quarter of the households still did not have PWA on their premises in the urban area, and these households were more likely to use groundwater, as well as relying only on groundwater for drinking and bathing purposes. The increased risk of HDO with the increasing *E. coli* concentration of the groundwater among households without PWA suggests the need for further studies on interventions to improve groundwater quality and their effect on reducing HDO. The interpolation method and interdisciplinary approach used in this study are relevant to other developing countries with similar settings.

Authors' contributions: SS performed the multivariate and geostatistical analysis, and drafted the manuscript. TN, EH and BM performed the groundwater field survey and data management in GIS. JM discussed and interpreted the surface prediction and interpolation via the geostatistical analysis. NK, YA and JS critically revised the manuscript for intellectual content. KN was involved in the groundwater field survey, microbial analysis, discussions on drafting the manuscript and project management. All authors read and approved the final manuscript. SS and KN are the guarantors of the paper.

Acknowledgments: The authors are grateful to the Asian Development Bank and Dr Kanako Yoden for providing data, and to Dr Salina Shrestha for helping with ethical issues. They would also like to acknowledge Mr Bijay M. Shakya and Mr Anil Aryal, PhD students of University of Yamanashi, for kindly providing information on hydrogeology and rainfall data.

Funding: This study was supported by the Global Centre of Excellence program and the Grant-in-Aids for Scientific Research [No. 25630226] of the JSPS. It was also partly supported by the Science and Technology Research Partnership for Sustainable Development Program (SATREPS, Project Manager: Professor Narendra Man Shakya and Professor Futaba Kazama) of the Japan International Cooperation Agency (JICA)/Japan Science and Technology Agency (JST).

Competing interests: None declared.

Ethical consideration: The ethical review board of the University of Yamanashi School of Medicine approved the study protocol.

References

1 Morris BL, Lawrence AR, Chilton P et al. Groundwater and its susceptibility to degradation: a global assessment of the problem and options for management. Nairobi, Kenya: United Nations Environment Programme, 2003.

2 IIED. Groundwater, self-supply and poor urban dwellers—a review with case studies of Bangalore and Lusaka. London: International Institute of Environment and Development, 2010.

3 Ferguson AS, Mailloux BJ, Ahmed KM et al. Hand-pumps as reservoirs for microbial contamination of well water. J Water Health 2011;9(4): 708–17.

4 Barthiban S, Lloyd BJ, Maier M. Applicability of a surveillance methodology for the microbiological safety of well water supplies, in a highly vulnerable hydrogeological setting. J Water Resour Prot 2012; 4(11):899–912.

5 Uy D, Haka S, Huya C et al. Comparison of tube-well and dug-well groundwater in the arsenic polluted areas in Cambodia. In: Fukusi K, Kurisu F, Oguma K et al (Eds). Southeast Asian Water Environment 4. London: International Water Association Publishing, 2010;41–46.

6 Krishnan RR, Dharmaraj K, Kumari B. A comparative study on the physicochemical and bacterial analysis of drinking, borewell and sewage water in the three different places of Sivakasi. J Environ Biol 2007;28(1):105–8.

7 Murphy H, Thomas M, Schmidt P et al. Estimating the burden of acute gastrointestinal illness due to Giardia, Cryptosporidium, Campylobacter, *E. coli* O157 and norovirus associated with private wells and small water systems in Canada. Epidemiol Infect 2016; 144(7):1355–70.

8 Richards RP, Baker DB, Creamer NL et al. Well water quality, well vulnerability, and agricultural contamination in the Midwestern United States. J Environ Qual 1996;25(3):389–402.

9 Schets FM, During M, Italiaander R et al. *Escherichia coli* O157: H7 in drinking water from private water supplies in The Netherlands. Water Res 2005;39(18):4485–93.

10 ADB. Kathmandu Valley Water Supply and Waste Water System Improvement Project. Mandaluyong: Asian Development Bank, 2010.

11 Shrestha S, Aihara Y, Bhattarai AP et al. Dynamics of domestic water consumption in the urban area of the Kathmandu Valley: situation analysis pre and post 2015 Gorkha Earthquake. Water 2017;9:222.

12 Shrestha S, Nakamura T, Malla R et al. Seasonal variation in the microbial quality of shallow groundwater in the Kathmandu Valley, Nepal. Water Sci Tech-W Sup 2014;14(3):390–7.

13 Haramoto E, Yamada K, Nishida K. Prevalence of protozoa, viruses, coliphages and indicator bacteria in groundwater and river water in the Kathmandu Nalley, Nepal. Trans R Soc Trop Med Hyg 2011;105:711–16.

14 IHME. GBD Profile. Nepal: Institute for Health Metrics and Evaluation, 2010.

15 United Nations Inter-agency Group for Child Mortality Estimation. Levels & trends in child mortality. Geneva: UN Inter-agency Group for Child Mortality Estimation.

16 Ashraf S, Yunus M. Waterborne diseases of bacterial origin in relation to quality of water in a suburb of Uttar Pradesh. Biomed Environ Sci 1997;10:442–50.

17 Shrestha S, Aihara Y, Yoden K et al. Access to improved water and its relationship with diarrhoea in Kathmandu Valley, Nepal: a cross-sectional study. BMJ Open 2013;3:e002264.

18 UNICEF/WHO. Progress on drinking water and sanitation: 2012 Update. https://www.unicef.org/media/files/JMPreport2012.pdf [accessed April 15, 2015].

19 Tumwine JK, Thompson J, Katua-Katua M et al. Diarrhoea and effects of different water sources, sanitation and hygiene behavior in East Africa. Trop Med Int Health 2002;7(9):750–6.

20 Yoshida M, Igarashi Y. Neogene to quaternary lacustrine sediments in the Kathmandu Valley, Nepal. J Nepal Geol Soc 1984;4:73–100.

21 Jha MG, Khadka MS, Shrestha MP et al. The Assessment of Groundwater Pollution in the Kathmandu Valley, Nepal. Report, Joint Australia (AGSO)—Nepal (GWRDB) Project, 1994–1996. Nepal: Australian Geological Survey Organization, Australia and Ground Water Resources Development Board, 1997.

22 Shrestha SD, Karmacharya R, Rao GK. Estimation of groundwater resources in Kathmandu Valley, Nepal from the water table fluctuation and the natural discharge. J Groundwater Hydrol 1996;38(1):29–40.

23 Warner NR, Levy J, Harpp K et al. Drinking water quality in Nepal's Kathmandu Valley: a survey and assessment of selected controlling site characteristics. Hydrogeol J 2008;16:321–34.

24 CBS. National Population Census 2011—Household and population by sex ward level. Thapathali Kathmandu: Central Bureau of Statistics, 2012.

25 KUKL. KUKL- At a Glance, Third Anniversary. Kathmandu: Kathmandu Upatyaka Khanepani Limited, 2010.

26 Ferguson AS, Layton AC, Mailoux BJ et al. Comparison of fecal indicators with pathogenic bacteria and rotavirus in groundwater. Sci Total Environ 2012;431:314–22.

27 Shrestha S, Shrestha S, Shindo J et al. Virological quality of irrigation water sources and pepper mild mottle virus and tobacco mosaic virus as index of pathogenic virus contamination level. Food Environ Virol 2018;10(1):107–20.

28 Owusu-Ansah, EGJ, Sampson A, Amponsah SK et al. Probabilistic quantitative microbial risk assessment model of norovirus from wastewater irrigated vegetables in Ghana using genome copies and fecal indicator ratio conversion for estimating exposure dose. Sci Total Environ 2017;601–2:1712–19.

29 Nas B, Berktay A. Groundwater quality mapping in urban groundwater using GIS. Environ Monit Assess 2010;160:215–27.

30 Burgess T, Webster R. Optimal interpolation and isarithmic mapping of soil properties. J Soil Sci 1980;31:333–41.

31 Shrestha S, Haramoto E, Malla R et al. Risk of diarrhoea from shallow groundwater contaminated with enteropathogens in of the Kathmandu Valley. J Water Health 2015;15:259–69.

32 Bohling G. Kriging. Kansas Geological survey. http://people.ku.edu/~gbohling/cpe940/Kriging.pdf [accessed March 24, 2018].

33 JICA. Groundwater management project in Kathmandu Valley, Final report, Main report and supporting reports. Kathmandu: Japan International Cooperation Agency, 1990.

34 Shrestha SD. Geology and hydrogeology of groundwater aquifers in Kathmandu Valley. In: Shrestha S, Pradhananga D, Pandey VP (Eds). Kathmandu Valley Groundwater Outlook. Nepal: Asian Institute of Technology, The Small Earth Nepal, Center of Research for Environment Energy and Water, International Research Centre for River Basin Environment - University of Yamanashi, 2012;21–30.

35 Murphy HM, Prioleau MD, Borchardt MA et al. Review: epidemiological evidence of groundwater contribution to global enteric disease, 1948–2015. Hydrogeol J 2017;25:981–1001.

36 Razzolini MT, Weir MH, Matter MH et al. Risk of Giardia infection for drinking water and bathing in a per-urban area in Sao Paulo, Brazil. Int J Environ Health Res 2011;21:222–34.

37 Schets FM, Schijven JF, de Roda Husman AM. Exposure assessment for swimmers in bathing waters and swimming pools. Water Res 2011;45:2392–400.

38 Papastergiou P, Mouchtouri V, Pinaka O et al. Elevated bathing-associated disease risks despite certified water quality: a cohort study. Int J Env Res Public Health 2012;9:1548–65.

39 Shrestha S, Malla SS, Aihara Y et al. Water quality at supply source and point of use in the Kathmandu Valley. J Water Environ Tech 2013;11(4):331–40.

40 Alam N, Henry FJ, Rahaman MM. Reporting errors in one-week diarrhoea recall surveys: experience from a prospective study in rural Bangladesh. Int J Epidemiol 1989;18(3):697–700.