

Environmental predictors of diarrhoeal infection for rural and urban communities in south India in children and adults

D. KATTULA¹, M. R. FRANCIS¹, A. KULINKINA², R. SARKAR¹,
V. R. MOHAN³, S. BABJI¹, H. D. WARD^{1,4}, G. KANG^{1,4}, V. BALRAJ³ AND
E. N. NAUMOVA^{1,2*}

¹ *Division Gastrointestinal Sciences, Christian Medical College, Vellore, India*

² *Department of Civil and Environmental Engineering, Tufts University School of Engineering, Medford, MA, USA*

³ *Department of Community Health, Christian Medical College, Vellore, India*

⁴ *Division of Geographic Medicine and Infectious Diseases, Tufts Medical Center, Boston, MA, USA*

*Received 24 April 2014; Final revision 29 October 2014; Accepted 28 November 2014;
first published online 18 February 2015*

SUMMARY

Diarrhoeal diseases are major causes of morbidity and mortality in developing countries. This longitudinal study aimed to identify controllable environmental drivers of intestinal infections amidst a highly contaminated drinking water supply in urban slums and villages of Vellore, Tamil Nadu in southern India. Three hundred households with children (<5 years) residing in two semi-urban slums and three villages were visited weekly for 12–18 months to monitor gastrointestinal morbidity. Households were surveyed at baseline to obtain information on environmental and behavioural factors relevant to diarrhoea. There were 258 diarrhoeal episodes during the follow-up period, resulting in an overall incidence rate of 0·12 episodes/person-year. Incidence and longitudinal prevalence rates of diarrhoea were twofold higher in the slums compared to rural communities ($P < 0\cdot0002$). Regardless of study site, diarrhoeal incidence was highest in infants (<1 year) at 1·07 episodes/person-year, and decreased gradually with increasing age. Increasing diarrhoeal rates were associated with presence of children (<5 years), domesticated animals and low socioeconomic status. In rural communities, open-field defecation was associated with diarrhoea in young children. This study demonstrates the contribution of site-specific environmental and behavioural factors in influencing endemic rates of urban and rural diarrhoea in a region with highly contaminated drinking water.

Key words: Diarrhoea, incidence rate, risk factors, rural, urban.

INTRODUCTION

Diarrhoeal disease is the most important outcome of water-related intestinal infections and one of

the leading causes of childhood morbidity and mortality, especially in infants [1]. Of the 0·71 million annual diarrhoeal deaths, 90% occur in children, mainly from developing countries [2]. The reported incidence of diarrhoea in low- and middle-income countries is 2·9 episodes/child-year [3]. In India, more than 2·3 million children die every year and about 334 000 of these are attributed to diarrhoeal diseases [4, 5].

* Author for correspondence: E. N. Naumova, Ph.D., Associate Dean for Research, Professor, Department of Civil and Environmental Engineering, Tufts University School of Engineering, 200 College Avenue, Medford, MA 02155, USA.
(Email: Elena.Naumova@tufts.edu)

The World Health Organization (WHO) estimates that by providing basic sanitation, safe drinking water and practising good hygiene, diarrhoea can be reduced by 26%. However, 40% of the world's 6 billion people have no acceptable means of sanitation and more than 1 billion people draw their water from unsafe sources [6]. In India, over 70% of the rural population do not have sanitary toilets and do not practice any method of water purification [7]. Open-field defecation is still a common practice in rural areas, leading to contamination of the water table through years of seepage and may also contaminate drinking water because the water supply pipes and sewage channels are in close proximity, often crossing each other at street junctions [8]. Further, intermittent water supply results in a negative pressure in pipelines and the ingress of sewage which, in turn, can result in diarrhoeal outbreaks [9, 10]. The presence of animals in close proximity to human dwellings adds to the risk of transmission of zoonotic infections directly to humans or through the contaminated water [11, 12].

In India, a total of 14.3 million people migrated from rural to urban areas between 1991 and 2001; these slum dwellers are exposed to additional health risks due to the increasing population density and poor living conditions [13]. Poverty and unsanitary living conditions are known to increase diarrhoeal risk in children especially those aged <5 years [14]. Cumulative exposure to these factors combined with the lack of adequate sanitation facilities and documented contamination of drinking water make urban populations more vulnerable to diarrhoeal diseases [15, 16]. Only a few studies have attempted to compare diarrhoeal disease burden and its contributors in rural and urban communities.

Pathogens causing diarrhoea have complex dynamic patterns of transmission influenced by changing person-to-person contacts and other time-varying endogenous and exogenous risk factors resulting in substantial differences in estimated infection rates for various populations. In order to develop sustainable mechanisms for preventing diarrhoeal disease, it is essential to understand environmental, socio-cultural and behavioural differences that can contribute to the different rates of infections in rural and urban settings. This study was conducted to provide a set of measures of diarrhoeal disease burden in an 18-month longitudinal follow-up of individuals and families, and to identify factors in the domestic environment that can predict diarrhoea in urban and rural communities. Age-specific incidences of

diarrhoeal episodes and their duration were quantified in children aged <5 years in 300 households of rural and urban communities in southern India. The effect of environmental factors related to water, sanitation and hygiene (WASH) on diarrhoeal incidence was also explored, controlling for socioeconomic status (SES).

METHODS

Study area and population

Following Institutional Review Board clearance at both partnering institutions [Christian Medical College (CMC), Vellore, India and Tufts University School of Medicine, Boston, MA, USA], three rural sites, namely A. Kattupadi (KP), Kattuputhur (KT) and K. Pudur (PT), were chosen from 82 villages in Kaniyambadi block [population: 104 792, Community Health and Development (CHAD), CMC Vellore unpublished census data, 2008], south of Vellore town (located 12.9202° N, 79.1333° E). An existing census showed that, there were 4304 persons in the three villages and their major income source was agricultural labour (CHAD unpublished census data). Water supply was from deep-bore wells piped into overhead tanks and then through a network of pipes and taps [8, 10]. Most houses have no usable toilet and hence open-field defecation is a common practice [17].

The two urban sites were Ramanaickanplayam (RNP) and Kaspas (KS), two geographically adjacent semi-urban slums (defined as a compact settlement of households that are crowded together, having poorly built tenements, mostly of temporary nature, and with inadequate drinking water and sanitary facilities [18]), which were selected from four geographically contiguous areas located on the western outskirts of Vellore town with a population of 24 843 (unpublished census data collected by the study staff). The majority of urban residents earned their wages through unskilled labour and 'beedi-work' (manual production of indigenous cigarette-like tobacco products). The primary source of drinking water to these sites was piped water, sourced from deep-bore wells in a dry river bed about 5 miles away and supplied through a network of pipes and taps intermittently (at intervals of 2–28 days) by the local municipality. During periods of water scarcity, water was supplied by the municipality through water tankers [19].

Earlier studies have shown extensive drinking water contamination in both urban and rural Vellore

[8, 19–21], and poor hygiene and lack of sanitation facilities [17, 19]. Water-related outbreaks have also been reported in the past [10, 22–25]. Moreover, even though chlorination with bleaching powder of the overhead tanks is the recommended method of disinfection, surveys in both outbreak and non-outbreak settings have failed to find residual chlorine in the water samples, suggestive of inadequate chlorination of drinking water [8, 10, 19]. The three rural and two urban sites were selected based on convenience and willingness to participate. These sites, however, are similar to each other and not different from other villages or urban slums in and around Vellore, in terms of availability of water and sanitation facilities, access to roads or healthcare, climatic conditions and weather patterns.

Enrolment and data collection

Prior to recruitment, the study area was enumerated by trained study staff in a door-to-door survey (census) to identify households with children aged <5 years. The entire household along with the youngest child (referred to as the index child) was recruited for 18 months of follow-up. Written informed consent was obtained from the head of the household prior to recruitment. Verbal consent was obtained from all members of the family, and assent was obtained from the children. Recruitment was systematic; until the required sample size was reached. The entire field staff were trained and periodically retrained in the use of questionnaires and case definitions; and their activities and the data collected every week were reviewed by two supervisors (D.K. and M.F.).

All recruited families were visited once every week and interviewed about diarrhoea, vomiting and other gastrointestinal symptoms experienced by any member of the family on each day since the last visit, using a structured questionnaire. If a family was not available, a phone interview was conducted and the period of stay outside their normal residence was noted. The case definition for a diarrhoeal episode was based on the WHO definition of ‘passage of 3 or more loose or watery stools in 24 hours or in case of infants, more frequent than normal passage of watery stool’ [26]. A new episode of diarrhoea was defined if it occurred at least 48 h after the end of the previous episode.

When a participant reported an episode of diarrhoea, a single stool sample was collected from the individual within 1 week of the end of the diarrhoeal episode. All the stool samples were collected in screw-

capped plastic containers and transferred to the Wellcome Trust Research Laboratory at CMC, within 4 h of collection in an ice-packed box. Stool containers (without preservative) were left with the families and phone numbers of the field workers were provided. Families were asked to contact the field staff if there was an episode of diarrhoea in the family.

Microscopy, ELISA and culture were performed to detect parasites (*Cryptosporidium*, *Giardia*, *Entamoeba histolytica*, *Hymenolepis nana*, *Ascaris*, *Trichuris*, hookworm) viruses (rotavirus genogroup A) and bacterial (*Shigella*, *Vibrio*, *Escherichia coli*) pathogens, respectively. If the culture was positive for *E. coli*, polymerase chain reaction (PCR) was performed to differentiate the pathotype. The detailed laboratory methods for the testing of viral, bacterial and parasitic pathogens in the diarrhoeal stool samples are presented as Supplementary material (Supplementary Document S1). Water samples from public taps and household containers were collected monthly and tested for total and faecal coliform counts (see Supplementary Table S2). The methods used for collection and testing of water samples and the water quality results are the subject of another paper (A. Kulinkina *et al.*, unpublished data), and are not presented in detail in the current paper.

Duration of diarrhoea was calculated as the difference between the start and end dates of each episode. Incidence rates were calculated as the number of episodes divided by the person-years of follow-up. Longitudinal prevalence [27] was calculated as total days of diarrhoea divided by the person-years of follow-up.

During their household visits, the field workers also obtained baseline demographic details of the participating households, and information on water treatment, usage and storage, sanitation, animal contact, and household hygiene practices. Households with ≥ 2 married siblings/cousins living together and sharing the same kitchen were classified as ‘joint’ families; the families were termed ‘extended’ if grandparents resided in the same household. Houses were also categorized by the type of roof and flooring: a house with thatched/tiled roof or earthen floor was termed a ‘kutchra’ house, whereas a house with a concrete roof and floor was called a ‘pucca’ or permanent house. Crowding was defined as >5 individuals living per room in a household, and SES was assessed using the Modified Kuppuswamy SES scale based on average household income per month, education and occupation of the head of a household [28].

Data analysis

Data were entered using Epi-Info 2002 (CDC, USA) software and analysed using Stata v. 10.1 for Windows (StataCorp, USA). Comparison of baseline differences between the rural and urban cohorts was tested using χ^2 test or Fisher's exact test for categorical variables and two-tailed *t* tests or Wilcoxon rank sum tests for continuous variables, including incidence and longitudinal prevalence rates. For the analysis, upper/upper middle SES strata were grouped as upper SES, lower middle as middle SES and upper lower/lower as low SES. Univariate comparisons were restricted to indicators with sufficient number of cases per category and the results are provided in Supplementary Tables S4 and S5. Indicators of hygiene practice with more than 95% positive responses, such as having a dedicated container for storing drinking water and covering the water container or food, were not included in the analysis.

Incidence and duration of diarrhoea were considered as separate health outcomes for the analysis due to heterogeneity of behavioural and environmental factors for each outcome. A generalized log-linear regression model (GLM) with a negative binomial distribution assumption for these health outcomes was used, adjusted for clustering effect at the community level. Separate models were developed for rural and urban areas, as well as for index children and all individuals. All environmental and demographic factors, identified in the univariate analysis as significant at $P < 0.05$ were considered for inclusion in the full multivariate models. A parsimonious regression model was chosen considering the significance of and correlation among predictors in the full model. For direct comparison of rural and urban communities, some non-significant variables were also retained in the final model. The results are presented as incidence risk ratios (IRR) along with 95% confidence intervals (CI), where the lower boundary of the CI greater than 1 typically refers to a significant effect at $\alpha = 0.05$. The quality of fit of each model is presented as the percentage of variability explained based on the difference between the null and residual deviances.

RESULTS

The survey identified 862 households (158 rural, 719 urban), which fulfilled the eligibility criteria (at least one child aged <5 years). Of these, based on geographical contiguity in the urban area, 300

(140 rural, 160 urban) households were recruited, comprising a total of 1579 individual participants (727 in rural and 852 in urban areas). A total of 279 households (93%, 133 rural, 146 urban) remained in the study from recruitment until the end of follow-up. Households were followed for a median period of 17.6 [interquartile range (IQR) 12.9–18.0] months for a total of 4790 household-months. Each individual was followed for a median of 17.6 (IQR 15.7–17.9) months for a total of 25 460 person-months. The primary reason for loss to follow-up was migration out of the study area, but 13 deaths were also reported, in which two were children aged <5 years. Households that eventually left the study contributed a median period of 8.2 (IQR 3.8–11.7) months of person-time. The cohort recruitment and follow-up is outlined in Figure 1.

Demographic and socioeconomic profiles

The overall median age at the time of recruitment of the family was 25 (IQR 5–37) years with children aged <5 years representing 24% of the recruited population (see Fig. 2). The majority of families in rural areas were Hindus (84%) followed by Christians (16%), whereas in urban areas 33% were Hindus and 66% were Muslims ($P < 0.0001$). In rural areas, 70% of households were joint families, compared to 54% nuclear households in the urban study areas ($P < 0.01$); a typical household, both in rural and urban areas, had five members.

Most (66%) houses in urban and rural areas had a concrete roof and floor. In rural areas, 82% of the families owned the house they resided in, while only 43% in urban areas lived in their own house. Firewood (56%) was the main cooking fuel in rural areas followed by liquefied petroleum gas (LPG) (40%), whereas in urban areas it was predominantly LPG (53%) followed by kerosene (31%). While the majority (78%) of households were from a low socioeconomic strata in both rural and urban areas, the average monthly household income in the rural households was higher than their urban counterparts (Rs. 4500 vs. Rs. 3000, $P = 0.003$).

Water, sanitation, and hygiene practices

Almost all households (98%) had dedicated container(s) for storing drinking water. In the rural areas fewer families (24%) purified water compared to the urban areas (66%). Half the rural families (58%) and

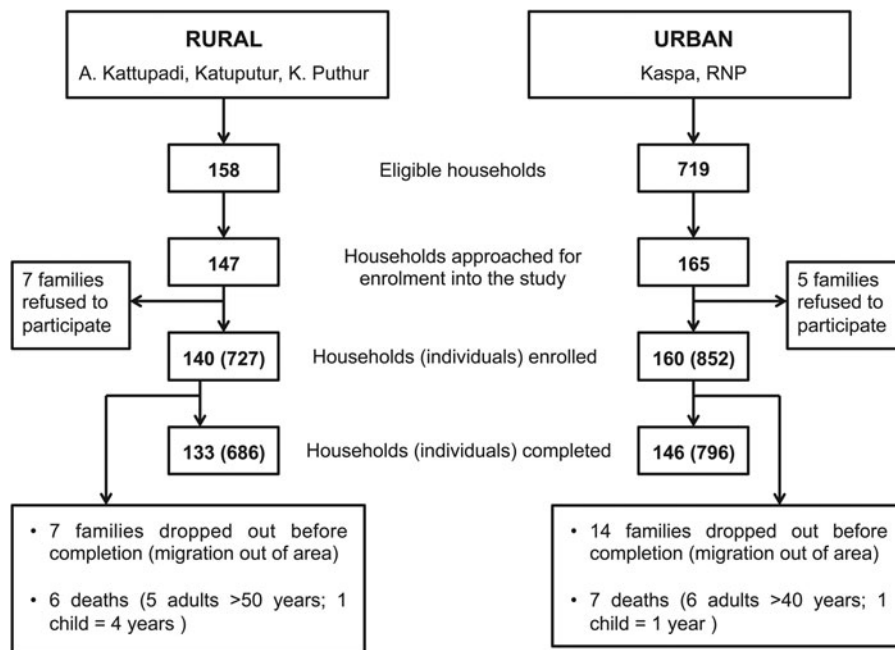


Fig. 1. Establishment of the cohort and follow-up.

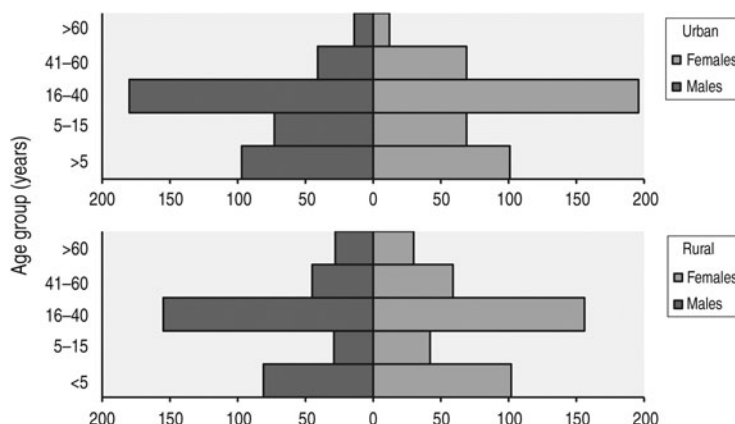


Fig. 2. Comparison of age distribution of the study population in urban and rural areas.

98% of urban families owned toilets; the practice of open-field defecation was common in rural areas with nearly 55% of the families reporting this practice. Domesticated animals within or close to households were reported in greater proportion by rural families (61%) compared to urban families (18%). A comparison of water and sanitation practices between the rural and urban participants is presented in Table 1. The majority of the household water samples tested (84% in rural and 85% in urban households) had >10 faecal coliform colonies/100 ml (see Supplementary Table S2).

Diarrhoeal episodes and age-specific incidence rates

Between August 2010 and March 2012, a total of 258 episodes of diarrhoea (74 in rural and 184 in urban areas) were reported in 49 rural and 105 urban residents. The overall incidence (95% CI) of diarrhoea was 0.12 (0.11–0.14) episodes/person-year of observation; the incidence was significantly higher in urban areas (0.15, 0.13–0.17) than in rural areas (0.08, 0.06–0.10, *P* < 0.01). The proportion of time with diarrhoea (average longitudinal prevalence) in urban areas was almost twofold higher than in rural

Table 1. Comparison of water usage and sanitation practices in the rural and urban areas of the study population

	Rural n (%)	Urban n (%)
Dedicated drinking water storage container	139 (99)	156 (98)
Covered water container	138 (99)	154 (96)
Purification of water		
Always	17 (12)	105 (66)**
Occasionally	17 (12)	17 (11)
Never	106 (76)	38 (24)
Method of purification†		
Filter cloth/sieve	7 (5)	102 (64)**
Filter using ceramic filter	3 (2)	2 (1)*
Boiling	28 (20)	37 (23)**
Packaged water	0 (0)	8 (5)
Never	106 (76)	38 (24)
Covering food	139 (99)	158 (99)
Disposal of household/kitchen waste		
In a corner within compound (backyard)	68 (49)	81 (51)
Just outside the compound	58 (41)	64 (40)
Designated garbage disposal areas/bins	38 (27)	110 (69)**
In dug out pits within/outside compound	32 (23)	1 (1)**
Burning	69 (49)	4 (3)**
Presence of latrine in/near house	81 (58)	156 (98)**
Water in the latrine		
Running water through tap	17 (12)	54 (34)
In a bucket	64 (46)	99 (62)*
Place of defecation		
Latrine	63 (45)	159 (99)
Open field	77 (55)	1 (1)**
Wash hands after defecation		
With soap and water	53 (38)	99 (62)**
Only water	87 (62)	61 (38)
Presence of domesticated animals in/near the house	86 (61)	28 (18)**
Presence of flies in and around house	134 (96)	123 (77)**
Presence of animal shed		
In close proximity to house	80 (57)	24 (15)
Not present in the neighbourhood	60 (43)	136 (85)**
Use of cow dung		
Use cow dung cakes/manure	126 (90)	6 (4)**
Not in contact with cow dung	14 (10)	154 (96)
Handle animal waste		
Bare hands	94 (67)	5 (3)
Broom	32 (23)	2 (1)

† Percentages add to more than 100% as some families used more than one water purification method.

* $P < 0.01$, ** $P < 0.0001$.

areas [0.35 (0.32–0.39) vs. 0.17 (0.14–0.19), $P < 0.01$, respectively]. In adults, diarrhoeal incidence increased with increasing age – from 0.01 (0.006–0.02) episodes/person-year in the 15–40 years age group, to 0.02 (0.01–0.05) in 41–60 years age group, and to 0.05 (0.02–0.11) in those aged ≥ 60 years. There were no differences in occurrence of diarrhoea between males and females ($P = 0.90$).

Two hundred and nineteen episodes (85%) of the 258 diarrhoeal episodes were contributed by

120 children aged < 5 years; 200 (91%) of these 219 episodes were in 105 of the youngest child in the families (referred to as the index child). The overall incidence of diarrhoea in children aged < 5 years was 0.51 (0.44–0.58) episodes/person-year (see Table 2). Children aged < 5 years in urban areas had higher incidence than those in rural areas [0.67 (0.57–0.78) vs. 0.33 (0.26–0.42) episodes/person-year, $P = 0.01$, respectively]. During the first year of life, the longitudinal prevalence rate of diarrhoea was fourfold higher in

Table 2. Description of diarrhoeal episodes, incidence and longitudinal prevalence rates in children aged <5 years in the study population

	Age groups			
	<12 months	12–24 months	25–59 months	<5 years
Number of children (at enrolment)	38	182	161	381
Number of episodes				
Overall	52	102	65	219
Rural*	22	39	5	66
Urban*	30	63	60	153
Incidence rate (per child per year)				
Overall	1.42 (1.08–1.86)	0.51 (0.42–0.62)	0.33 (0.26–0.43)	0.51 (0.44–0.58)
Rural*	0.96 (0.64–1.50)	0.35 (0.25–0.48)	0.07 (0.03–0.18)	0.33 (0.26–0.42)
Urban*	2.20 (1.50–3.10)	0.71 (0.56–0.92)	0.48 (0.36–0.61)	0.67 (0.57–0.78)
Longitudinal prevalence rate (per child per year)				
Overall	5.88 (5.15–6.73)	1.75 (1.58–1.95)	1.00 (0.87–1.15)	1.77 (1.65–1.89)
Rural*	2.59 (2.00–3.34)	1.31 (1.11–1.54)	0.28 (0.18–0.44)	1.11 (0.97–1.27)
Urban*	11.00 (9.69–13.01)	2.32 (2.02–2.66)	1.38 (1.19–1.60)	2.34 (2.15–2.55)

* The difference between the outcomes measured in rural and urban areas was significant ($P < 0.01$, two-tailed Wilcoxon rank sum tests) for all groups.

Values in parentheses are 95% confidence intervals.

urban slums than in rural areas and the incidence of diarrhoea in urban slums was nearly twofold higher compared to the rural area. Both outcomes declined rapidly with age.

Both urban and rural households had a median duration of diarrhoea of 2 days. Eighty-four (33%) diarrhoeal events were associated with either fever or vomiting, and 43 (17%) episodes were associated with both fever and vomiting; there was no rural–urban difference in the proportion of diarrhoeal episodes with concomitant symptoms ($P = 0.80$).

Pathogen specific stool sample positivity

A total of 209 (51 rural, 158 urban) stool samples were collected from the 258 reported cases of diarrhoea in the study participants. Forty-nine (19%) reported cases of diarrhoea did not have a laboratory test result because these samples were not provided within the 7-day window period, or due to non-availability of the participant from the study area during the diarrhoeal episode. Pathogens were isolated from 85 (41%) (19 rural, 66 urban, $P = 0.56$) stool samples; of which 64 (75%; 15 rural, 49 urban) samples had a single pathogen and 21 (25%; 4 rural, 17 urban) had multiple pathogens; the difference was not statistically significant ($P = 0.67$). The pathogens isolated from the diarrhoeal samples and their age-wise distribution are given in Table 3 and Supplementary Table S3,

Table 3. Description of diarrhoeal episodes and pathogen distribution (%) in rural and urban areas of the study population

Pathogen	Overall n (%)	Rural n (%)	Urban n (%)	<i>P</i> value
Number of samples	209 (100)	51 (24)	158 (76)	–
<i>Giardia</i>	27 (13)	6 (12)	21 (13)	0.77
<i>Cryptosporidium</i>	6 (3)	0 (0)	6 (4)	0.15
<i>Ascaris</i>	3 (1)	0 (0)	3 (2)	0.32
Rotavirus	9 (4)	3 (6)	6 (4)	0.52
<i>Shigella</i>	8 (4)	2 (4)	6 (4)	0.96
<i>Vibrio</i>	1 (0.5)	1 (2)	0 (0)	0.07
<i>Hymenolepis nana</i>	2 (1)	0 (0)	2 (1)	0.41
<i>Entamoeba histolytica</i>	1 (0.5)	1 (2)	0 (0)	0.07
Total diarrhoeagenic <i>E. coli</i> *	49 (23)	10 (20)	39 (25)	0.45
EAEC	21 (10)	5 (10)	16 (10)	0.94
EPEC	13 (6)	4 (8)	9 (6)	0.58
EIEC	5 (2)	0 (0)	6 (4)	0.15
EPEC	14 (6.7)	3 (5.9)	11 (7.0)	0.78

EAEC, Enteraggregative *E. coli*; ETEC, enterotoxigenic *E. coli*; EIEC, enteroinvasive *E. coli*; EPEC, enteropathogenic *E. coli*.

* More than one *E. coli* pathotype identified in four samples, hence the sum of individual pathotypes is greater than the total of number of diarrhoeagenic *E. coli* positive samples.

respectively. Diarrhoeagenic *E. coli* was detected in 49 (23%) and *Giardia* identified in 27 (13%) samples. All samples positive for *Ascaris* spp. and *Cryptosporidium* spp. were from the urban study areas. *Shigella* spp. was detected in eight (4%) samples, of which 75% were from urban and 25% were from rural diarrhoeal samples. A total of 17 (20%) of the 85 positive samples had more than one pathogen isolated (two pathogens in 13 samples and three pathogens in four samples). Of the 49 diarrhoeagenic *E. coli*-positive samples, more than one *E. coli* pathotype (all involving enteroaggregative *E. coli*) was identified in four (23%) samples.

Risk factors associated with diarrhoeal episodes and duration of diarrhoea at the household level

Univariate analysis of potential risk factors indicated that the presence of young siblings (<5 years) in a household was a significant risk factor for diarrhoea in urban (values given are IRR, 95% CI) (1.95, 1.52–2.50) and rural (3.46, 2.72–4.41) areas. Crowding was also a risk factor in both areas [2.11 (1.15–3.89) and 1.31 (0.21–2.02) in rural and urban areas, respectively]. A similar association was observed for the longitudinal prevalence of diarrhoea (see Supplementary Table S4).

In the multivariate analysis for the rural area, the presence of young siblings was associated with an increase in the incidence of diarrhoea at the household level (3.15, 2.40–4.13), while the use of latrines offered significant protection against diarrhoea (0.55, 0.42–0.71). In the urban area, presence of young siblings (1.92, 1.62–2.27), low SES (1.27, 1.02–1.59), and washing fruits and vegetables with drinking water (1.55, 1.33–1.81) were all identified as independent risk factors.

In the multivariate analysis, the longitudinal prevalence of diarrhoea in the rural areas increased in households with young siblings (3.09, 1.91–5.00), crowded living conditions (1.83, 1.22–2.75) and houses with domesticated animals in close proximity (2.39, 0.82–6.95). The longitudinal prevalence of diarrhoea in urban areas also increased in households with young siblings (2.54, 1.61–4.01), crowded settings (1.36, 1.04–1.79) and in households with low SES (1.70, 1.34–2.17).

Risk factors associated with diarrhoeal episodes and duration of diarrhoea in children aged <5 years

Univariate analysis (Supplementary Table S5) demonstrated that, in rural areas, the incidence of diarrhoea

significantly increased in the youngest of the children compared to their siblings aged <5 years (IRR, 95% CI) (3.17, 2.19–4.57), and in households practising open-field defecation (2.38, 1.58–3.58). The presence of a latrine in the house offered protection against diarrhoeal episodes (0.49, 0.38–0.64). In urban areas, the risk of diarrhoea increased for index children living with siblings aged <5 years (2.02, 1.16–3.15), in crowded households (1.36, 0.98–1.87) and households with low SES (1.24, 1.01–1.51). The use of boiled drinking water (0.86, 0.85–0.87) and the presence of a continuous supply of water in latrines (0.62, 0.40–0.97) were identified as protective factors. Presence of domesticated animals in close proximity to the home (0.56, 0.40–0.77), or animal sheds in close proximity to the household (0.52, 0.44–0.62), were also identified as protective factors.

Table 4 shows the results of the multivariate analysis for incidence and longitudinal prevalence of diarrhoea for index children. In both rural and urban areas, presence of young siblings was a risk factor. In the rural areas the presence of domesticated animals in close proximity to the home (2.80, 2.05–3.79) and defecation in an open field (1.53, 1.00–2.36) were associated with a significant increase in incidence and duration of diarrhoea [3.97 (3.34–4.72), and 1.45 (1.07–1.96), respectively].

DISCUSSION

To our knowledge, this is the first longitudinal study in India that estimates rural–urban differences in the magnitude of diarrhoeal disease burden, and attempts to assess risk factors associated with these differences.

Acute diarrhoeal outbreaks have been commonly attributed to contaminated water worldwide [29]. In southern India, water in rural and urban areas consistently have high levels of contamination with coliforms and faecal coliforms [8, 10, 19]. Although the analysis of water samples collected from taps (source) as well as households (point-of-use) at the study sites showed evidence of widespread contamination of the drinking water, both in rural and urban areas (A. Kulinkina *et al.*, unpublished data), the diarrhoeal incidence in rural areas was far less than in urban areas suggesting that water may not be the primary mode of transmission of diarrhoeal illnesses in endemic settings.

A systematic review showed that the median incidence of diarrhoea in children aged <5 years in developing countries was 2.9 episodes/child-year, with the highest incidence observed in younger children

Table 4. *Multivariate regression analysis for risk factors for diarrhoea and duration of diarrhoea in the index children in rural and urban areas*

	Diarrhoeal episodes		Duration of diarrhoea	
	IRR (95% CI)	P value	IRR (95% CI)	P value
Rural				
Person-to-person transmission				
Presence of siblings aged <5 years	2.80 (1.88–4.13)	<0.001	2.62 (1.25–5.45)	0.01
Crowding (>5 individuals per room)	0.95 (0.50–1.81)	0.89	0.88 (0.63–1.20)	0.41
Socioeconomic indicators				
Kutch house	2.48 (0.98–6.23)	0.05	2.63 (1.04–6.64)	0.04
WASH indicators				
Presence of domesticated animal in the house	2.80 (2.05–3.79)	<0.0001	3.97 (3.34–4.72)	<0.0001
Animal shed in proximity to the household	0.99 (0.58–1.67)	0.98	0.85 (0.56–1.31)	0.48
Use of cow dung in household	0.62 (0.16–2.42)	0.49	0.35 (0.07–1.76)	0.20
Open field defecation	1.53 (1.00–2.36)	0.05	1.45 (1.07–1.96)	0.01
Deviance explained (%)	26.69		22.47	
Urban				
Person-to-person transmission				
Presence of siblings aged <5 years	1.97 (1.07–3.61)	0.03	2.85 (1.54–5.26)	0.001
Crowding (>5 individuals per room)	1.17 (0.58–2.34)	0.66	1.37 (0.75–2.50)	0.30
Presence of relatives >60 years	1.16 (0.44–3.07)	0.76	–	
Socioeconomic indicators				
Low socioeconomic status	0.98 (0.67–1.43)	0.93	1.24 (1.13–1.35)	<0.001
WASH indicators				
Presence of domesticated animal in the house	0.63 (0.44–0.90)	0.01	0.78 (0.48–1.28)	0.33
Boiling drinking water	1.05 (0.97–1.15)	0.19	1.14 (0.75–1.75)	0.53
Continuous water in latrine	0.65 (0.42–1.00)	0.05	0.74 (0.37–1.48)	0.39
Deviance explained (%)	7.54		9.79	

IRR, Incidence risk ratio; CI, confidence interval; WASH, water, sanitation and hygiene.

(4.5 episodes/child per year in children aged 6–11 months, falling progressively to 2.3 episodes/child per year by the age of 4 years) [3]. In our study, the average incidence of diarrhoea was highest in infants and decreased as the child grew older. Another longitudinal study from the same area conducted during 2002–2006 showed that the incidence of diarrhoeal illness declined from 3.6/child-year in the first year to 1.2 in the third year indicating decline of diarrhoeal rates with increase in age [30]. The incidence of diarrhoea in adults was low, 0.05 (95% CI 0.02–0.11) episodes/person-year. A study from Kenya also reported a similar trend of lower rates of diarrhoea for children aged >5 years [31], indicating that the burden of diarrhoea is higher in young children compared to older individuals.

This study showed a much lower diarrhoeal rate in children, but was consistent over the 18 months of follow-up. It also showed a similar trend in declining diarrhoeal incidence with age as with earlier studies. Other studies have found that a longer gap between

two study visits reduces the true disease rates due to recall bias, especially if the disease is mild and of short duration. A study from southern India showed a 45% decrease in reporting of diarrhoea in children when the recall period was beyond 3 days [32]. Another study from The Gambia also showed a 50% decrease in reporting of symptoms of diarrhoea for a recall period of 8 days [33]. Feikin *et al.* suggested not exceeding a recall period beyond 3 and 4 days in children and adults, respectively, to capture at least 80% of disease, particularly for longitudinal studies with repeated visits [31]. Although this study had weekly surveillance visits, some degree of under-reporting particularly of mild diarrhoeas could have occurred.

The present study consistently identified having a sibling aged <5 years in a household (both in rural and urban areas) as a significant risk factor, resulting in a higher incidence as well as a longer duration of diarrhoea. This is in agreement with previous studies where households with a high proportion of younger

children have been shown to be at a higher risk of having diarrhoeal disease [34–36]. In India, it is not uncommon for an older child to look after a younger sibling.

In our study, rural households with >5 individuals per room had longer duration of diarrhoea than those with a lower density of individuals. In the urban areas, crowding increased both the duration and incidence of diarrhoea. This finding is in contrast to a study from Jamaica, which did not show any association between the degree of crowding in households and diarrhoea [37]. Crowding in the household (higher density of people per room) increases person-to-person contact, thereby facilitating the spread of infection [38, 39], especially in areas with poor sanitation and hygiene.

Our study showed that households in the rural areas with domesticated animals in close proximity had an increased risk of diarrhoea. Among index children in rural areas, the presence of a domesticated animal in or near a household increased the likelihood of diarrhoea twofold and the duration fourfold, whereas in urban households, it exhibited a protective effect. This could be because of differences in animal handling practices between rural and urban communities. Zoonotic transmission can occur through a variety of means such as working closely with livestock, household pets, and soil or water being contaminated with animal faeces. The presence of an animal in or near the house increases contamination of the yard with animal faeces, thereby increasing the probability of diarrhoea in children coming into contact with it while playing [12]. A study by Engleberg *et al.* reported that dog ownership was associated with rotavirus infection [40], and another study from Iran showed that the presence of animals within or near the home increased the risk of hospitalization with diarrhoea for children [41]. It is also likely that households with and without animals have different household hygiene practices.

A recent study reported that even if only a few individuals in the household used a latrine, contamination of the immediate environment and farmlands surrounding the household was less intense than around households without a functional latrine [42]. Our study showed that in rural areas, usage of a latrine for defecation offered 41% protection against diarrhoea. Index children who defecated in the open field had a 50% higher chance of diarrhoea than children who defecated in a latrine. Place of defecation was not considered for analysis in the urban areas because all the households reported using latrines.

Contrary to expectation, our study found that children from urban households where fruits and vegetables were washed with their drinking water had a higher chance of developing diarrhoea. While this may be because the drinking water in this area was found to have high levels of faecal contamination [19], in a situation where water is scarce, inadequate amount of water usage for washing fruits and vegetables and/or reuse of water may also support this counter-intuitive finding.

A possible limitation of this study was that no transport media was used during transfer of stool specimens to the laboratory, although they were transported within 4 h of collection. This could have resulted in lower isolation of some bacterial pathogens. Further, as no stool sample was collected from persons without diarrhoea, this study could not establish a causal relationship between the pathogen and diarrhoea. Another limitation of this study was the unavailability of data on breastfeeding and child weaning practices. Even though exclusive breastfeeding has been shown to have a protective effect against diarrhoeal disease [43], the higher than expected diarrhoeal rates observed in infants is possibly due to the practice of early introduction of supplementary feeding in the study area [44, 45] which, in turn, exposes children to a contaminated environment at a much younger age, thereby negating any protective effect conferred due to breastfeeding. Although the differences in breastfeeding and child weaning practices in the two communities might have contributed to the differences in diarrhoeal incidence between rural and urban communities, a detailed exploration of environmental risk factors in infants should be considered. Further analysis of data collected in this study should also explore the effects of weather, climate and water contamination on diarrhoea.

CONCLUSIONS

This longitudinal follow-up of families with children aged <5 years was done in urban and rural areas known to have high levels of water contamination [8, 10, 19]. This study demonstrated a diarrhoeal burden that was double in the urban slums compared to the rural villages. Risk of diarrhoea was related to multiple, potentially highly interdependent factors including the presence of siblings, overcrowding, open-field defecation, etc., emphasizing that water, while very important, is unlikely to be the only risk factor for diarrhoeal transmission in endemic settings.

Regardless of study site, the diarrhoeal incidence was highest in infants followed by children aged <5 years, and much less in older age groups, indicating a decreasing trend of diarrhoea with age. Hence, effective planning of reinforced strategies targeting high-risk groups and improvement in personal and domestic hygiene is essential to reduce the burden of gastrointestinal illnesses in communities with high environmental contamination.

SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit <http://dx.doi.org/10.1017/S0950268814003562>.

ACKNOWLEDGEMENTS

The authors thank Mr Senthil, Mr Arul Raj, Mrs Bhuvaneshwari, Ms. Suganthi, Mr Srinivasan and Mr Karuna for their help in data collection and entry. The authors also thank the team at Wellcome Research Laboratory for conducting the laboratory testing. We are also indebted to the families of the urban slums of Ramnaickapalayam and Kaspaa and the families of rural villages of A. Kattupadi, Katuputur and K. Puthur for their participation and support.

This study was supported by the Indo US CDC/ICMR grant (IRIS ID 2008–07980). D.K., R.S. and V.M.R. were supported by training grant D43 TW0007392 (G.K.). A.K. was partly supported by the Fellowship of the Tufts Institute of the Environment.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Kosek M, Bern C, Guerrant RL.** The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bulletin of the World Health Organization* 2003; **81**: 197–204.
2. **Walker CL, et al.** Global burden of childhood pneumonia and diarrhoea. *Lancet* 2013; **381**: 1405–1416.
3. **Fischer Walker CL, et al.** Diarrhea incidence in low- and middle-income countries in 1990 and 2010: a systematic review. *BMC Public Health* 2012; **12**: 220.
4. **Bassani DG, et al.** Causes of neonatal and child mortality in India: a nationally representative mortality survey. *Lancet* 2010; **376**: 1853–1860.
5. **Black RE, et al.** Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* 2010; **375**: 1969–1987.
6. **WHO.** Meeting the MDG Drinking Water and Sanitation target: the urban and rural challenge of the decade. Geneva and New York: WHO/United Nations Children's Fund, 2006.
7. **Indian Institute of Population Survey (IIPS) and Macro International.** *National Family Health Survey (NFHS-3), 2005–06: India. Volume 1.* Mumbai: IIPS, 2007.
8. **Gopal S, et al.** 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* 2009; **129**: 233–241.
9. **Bora D, et al.** V. cholerae 01 outbreak in remote villages of Shimla district, Himachal Pradesh, 1994. *Journal of Communicable Diseases* 1997; **29**: 121–125.
10. **Sarkar R, et al.** Epidemiological investigation of an outbreak of acute diarrhoeal disease using geographic information systems. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2007; **101**: 587–593.
11. **Kelly P, et al.** *Environmental and Behavioral Risk Factors for Diarrheal Diseases in Childhood: A Survey in Two Towns in Morocco.* Washington D.C.: Bureau for Global Programs, USAID, 1999.
12. **Shivoga WA, Moturi WN.** Geophagia as a risk factor for diarrhoea. *Journal of Infection in Developing Countries* 2009; **3**: 94–98.
13. **Madhiwalla N.** Healthcare in urban slums in India. *National Medical Journal of India* 2007; **20**: 113–114.
14. **Melo MC, et al.** Incidence of diarrhea in children living in urban slums in Salvador, Brazil. *Brazilian Journal of Infectious Diseases* 2008; **12**: 89–93.
15. **McPike J, Luke N.** Rethinking the urban advantage: differences in child diarrhea across rural, urban non-slum and urban slum locations in India. *Population Association of America 2012 Annual Meeting.* San Francisco: Population Association of America, 2012.
16. **UNICEF.** *The State of the World's Children, 2012: Children in an Urban World.* New York: United Nations Children's Fund, 2012.
17. **Banda K, et al.** Water handling, sanitation and defecation practices in rural southern India: a knowledge, attitudes and practices study. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2007; **101**: 1124–1130.
18. **Government of India.** Report of the Pronab Sen Committee on Slum Statistics/Census. New Delhi: Ministry of Housing and Urban Poverty Alleviation, Government of India, 2010.
19. **Brick T, et al.** Water contamination in urban south India: household storage practices and their implications for water safety and enteric infections. *International Journal of Hygiene and Environmental Health* 2004; **207**: 473–480.
20. **Firth J, et al.** Point-of-use interventions to decrease contamination of drinking water: a randomized, controlled pilot study on efficacy, effectiveness, and acceptability

- of closed containers, *Moringa oleifera*, and in-home chlorination in rural South India. *American Journal of Tropical Medicine and Hygiene* 2010; **82**: 759–765.
21. **Rose A, et al.** Solar disinfection of water for diarrhoeal prevention in southern India. *Archives of Disease in Childhood* 2006; **91**: 139–141.
 22. **Kang G, et al.** Epidemiological and laboratory investigations of outbreaks of diarrhoea in rural South India: implications for control of disease. *Epidemiology and Infection* 2001; **127**: 107–112.
 23. **Pai M, et al.** An epidemic of diarrhoea in south India caused by enteroaggregative *Escherichia coli*. *Indian Journal of Medical Research* 1997; **106**: 7–12.
 24. **Ramakrishna BS, et al.** Isolation of *Vibrio cholerae* O139 from the drinking water supply during an epidemic of cholera. *Tropical Medicine and International Health* 1996; **1**: 854–858.
 25. **Sowmyanarayanan TV, et al.** Investigation of a hepatitis A outbreak in children in an urban slum in Vellore, Tamil Nadu, using geographic information systems. *Indian Journal of Medical Research* 2008; **128**: 32–37.
 26. **WHO.** *The Treatment of Diarrhoea: A Manual for Physicians and other Senior Health Workers*. Geneva: World Health Organization, 2005.
 27. **Morris SS, et al.** Is prevalence of diarrhea a better predictor of subsequent mortality and weight gain than diarrhea incidence? *American Journal of Epidemiology* 1996; **144**: 582–588.
 28. **Kumar N, et al.** Kuppaswamy's socioeconomic status scale-updating for 2007. *Indian Journal of Pediatrics* 2007; **74**: 1131–1132.
 29. **WHO.** *Global Water Supply and Sanitation Assessment 2000 Report*. Geneva: World Health Organization, 2000.
 30. **Gladstone BP, et al.** Burden of illness in the first 3 years of life in an Indian slum. *Journal of Tropical Pediatrics* 2010; **56**: 221–226.
 31. **Feikin DR, et al.** Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. *International Journal of Epidemiology* 2010; **39**: 450–458.
 32. **Ramakrishnan R, et al.** Influence of recall period on estimates of diarrhoea morbidity in infants in rural Tamilnadu. *Indian Journal of Public Health* 1999; **43**: 136–139.
 33. **Byass P, Hanlon PW.** Daily morbidity records: recall and reliability. *International Journal of Epidemiology* 1994; **23**: 757–763.
 34. **Perry S, et al.** Household transmission of gastroenteritis. *Emerging Infectious Diseases* 2005; **11**: 1093–1096.
 35. **Hung BV,** The most common causes of and risk factors for diarrhea among children less than five years of age admitted to Dong Anh Hospital, Hanoi, Northern Vietnam (dissertation). Oslo, Norway: University of Oslo, 2006.
 36. **Werber D, et al.** Preventing household transmission of Shiga toxin-producing *Escherichia coli* O157 infection: promptly separating siblings might be the key. *Clinical Infectious Diseases* 2008; **46**: 1189–1196.
 37. **Bachrach LR, Gardner JM.** Caregiver knowledge, attitudes, and practices regarding childhood diarrhea and dehydration in Kingston, Jamaica. *Revista Panamericana de Salud Pública* 2002; **12**: 37–44.
 38. **El-Gilany AH, Hammad S.** Epidemiology of diarrhoeal diseases among children under age 5 years in Dakahlia, Egypt. *Eastern Mediterranean Health Journal* 2005; **11**: 762–775.
 39. **Saha CK.** Dynamics of risk determinants regarding diarrhea affected slum children in urban Dhaka: a dysfunctional health care system. *Current Research Journal of Social Sciences* 2012; **4**: 304–313.
 40. **Engleberg NC, et al.** Epidemiology of diarrhea due to rotavirus on an Indian reservation: risk factors in the home environment. *Journal of Infectious Diseases* 1982; **145**: 894–898.
 41. **Khalili B, et al.** Risk factors for hospitalization of children with diarrhea in Shahrekord, Iran *Iranian Journal of Clinical Infectious Diseases* 2006; **1**: 131–136.
 42. **Ziegelbauer K, et al.** Effect of sanitation on soil-transmitted helminth infection: systematic review and meta-analysis. *PLoS Medicine* 2012; **9**: e1001162.
 43. **Lamberti LM, et al.** Breastfeeding and the risk for diarrhea morbidity and mortality. *BMC Public Health* 2011; **11**: S15.
 44. **Gladstone BP, et al.** Infant morbidity in an Indian slum birth cohort. *Archives of Disease in Childhood* 2008; **93**: 479–484.
 45. **Sarkar R, et al.** Burden of childhood diseases and malnutrition in a semi-urban slum in southern India. *BMC Public Health* 2013; **13**: 87.