**Distribution Agreement**



Signature:

______________________________ ______________

Philip Dollard Date

Diarrhea and the urban environment in India: a cross-section and longitudinal analysis

By

Philip Dollard
Master of Public Health

Epidemiology

Benjamin Lopman, MSc, PhD
Committee Chair

Diarrhea and the urban environment in India: a cross-section and longitudinal analysis

By

Philip Dollard

BA
Emory University
2013

Thesis Committee Chair: Benjamin Lopman, MSc, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2018

## Abstract

Diarrhea and the urban environment in India: a cross-section and longitudinal analysis

By Philip Dollard

India is home to almost one-fifth of the world's population and a third of the world's poor. There is a persistent health gap between rural populations, which typically have worse health outcomes, and rapidly growing urban centers, though the growth of urban slums has complicated the picture. Diarrheal illness is one of the world's leading cause of death and disability and is a major health concern in India. This paper uses data from two waves of the India Human Development Survey, IHDS-I and IHDS-II, to examine prevalence of diarrhea in urban and rural areas of India, as well as associated risk factors. Logistic regression is used for each separate survey wave, and generalized estimating equations are used to conduct longitudinal analysis on data from both survey waves combined. Univariate analysis showed urban environment to have a protective effect on diarrhea in IHDS-I (OR = 0.63, 95% CI: 0.59, 0.67), IHDS-II (OR = 0.78, 95% CI: 0.72, 0.84), and longitudinal analysis (OR = 0.71, 95% CI: 0.67, 0.75). However, in all three cases, the association became null when adjusted for household income, literacy, and piped water, suggesting that these are major contributing factors in the difference between the health of urban and rural populations in India. This paper uses longitudinal data to pinpoint important factors in the public health gap between urban and rural areas in a developing nation and adds to a growing body of literature implicating income inequality as having a negative effect on the health of nations, especially developing nations.

Diarrhea and the urban environment in India: a cross-section and longitudinal analysis

By

Philip Dollard

BA
Emory University
2013

Thesis Committee Chair: Benjamin Lopman, MSc, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2018

Table of Contents

# Chapter I

## Introduction

India is a country in the midst of a great transition. The country currently boasts a fifth of the world's population and one of the world's fastest-growing economies (1). As it looks to the future, however, India remains plagued by widespread poverty, corruption, and disease. According to a recent report from the World Bank, India is home to fully one third of the world's people that live in poverty (2). Reform efforts focused on bringing India into the new millennium are underway in areas such as agriculture and infrastructure, but still there exist many problems for which viable long-term solutions have yet to be found.

### India's Disease Burden

Communicable, nutritional, maternal, and neonatal diseases are and for some time have been a serious issue in India, which suffers from a higher burden of these diseases than other countries with similar economic indicators (3). As of 2016, three of the five most common causes of death in India were due to communicable disease: diarrheal disease, tuberculosis, and respiratory infections, were together responsible for 17.5% of all deaths, behind only ischemic heart disease and chronic obstructive pulmonary disorder in overall mortality (4). Diarrhea has been strongly associated with short-term conditions such as dehydration and weight loss, and a small but growing body of evidence shows an association between repeated bouts of diarrhea and low weight, malnutrition, and even cognitive deficits (5-7)

Diarrhea in particular is a serious issue. Within India, it is the infectious disease most responsible for death and disability, and the third most common contributor to death and disability overall. In 2016, it caused an estimated 777,991 deaths, and is more likely to affect vulnerable members of the population. Diarrheal disease is caused by many different pathogens, complicating prevention strategies; in India, the three most common causes are rotavirus, shigella, and campylobacter (8). All three are mainly spread via the fecal-oral route, in which the pathogen is shed in an infected person's stool and ingested by another person via contaminated food, water, or other objects (9-11). The most common risk factors for diarrhea therefore stem from unclean water and unsafe sanitation such as living in a house with multiple children, drinking untreated water, and regular contact with farm animals (12-14).

Though diarrhea continues to impact India, mortality from diarrhea has been steadily decreasing (15). The decreases, in India and the rest of the world, can largely be attributed to improvements in access to safe water, sanitation, and nutrition (8). The Global Burden of Disease project estimates that, between 2005 and 2015, worldwide deaths from diarrhea among all age groups have decreased by 21%, and deaths among children under 5 have decreased by 34%. In India, the decline has been even sharper: deaths from diarrhea have decreased by 60% overall, and by 88% in children under five (3). Such improvements are reflective of a broader trend known as the epidemiologic transition, in which the proportion of the disease burden from non-communicable diseases grows larger than the disease burden from communicable, neonatal, nutritional, and maternal disease (16). The epidemiologic transition occurs when countries become sufficiently advanced to address a majority of immediately preventable diseases, and is generally considered an indicator of advancement. In India, the proportion of disease burden from non-communicable diseases overtook the proportion of disease burden from communicable, nutritional, maternal, and neonatal diseases for the first time in 2003, though the transition in individual states occurred over a span of 24 years, from 1986 to 2010 (3). The reduction in morbidity from communicable diseases, and diarrhea in particular, can in part be attributed to improved sanitation by providing safer latrines and increased access to clean drinking water in impoverished areas (3). For example, India's Prime Minister, Naredra Modi, launched the "Clean India" campaign in 2014, which in part aims to eliminate the common practice of open defecation by the year 2019, as well as install 75 million toilets across the country (17). The project's success to date is debatable (18), but it is part of a larger pattern, in which India's burden of communicable diseases has decreased as the country becomes more developed and works to improve sanitation.

Though the decline in diarrhea deaths has been precipitous, the decline in incidence of diarrheal disease has not been nearly as steep (8). This is likely because case management, such as improved healthcare access and increasing use of oral rehydration solution, has improved at a greater pace than prevention. This also illustrates that, though it is relatively easy to treat, diarrhea is very difficult to prevent, and will continue to be a major issue for years to come.

## Urbanization

Another facet of India's transition is its urbanization, the overall increase in the proportion of a population living in urban rather than rural areas. This decade marks the first time in history that the majority of humans are living in urban environments: as of 2014, the UN reported that 54% of the world's population was living in cities. As populations increase and the global economy shifts away from agriculture towards mass industry and technology, the proportion of people living in cities will continue to increase, with almost all population growth occurring in developing countries. Asia and Africa, for example, are both expected to increase their urban populations by roughly 20 percent by 2050, while the populations of over half of European cities are expected to decline in the same time frame (19). It is also important to note that urban environments can include anything from mega-cities with over ten million inhabitants to trading centers and towns with only a few thousand people. In India, about 377 million people, or roughly one third of the population, live in urban centers (20), and by 2050 that proportion is expected to grow to half the population (21), with half of the projected urban growth predicted to be in population centers with less than fifty thousand inhabitants. The number of urban centers has also increased dramatically, driven by industrialization and an increase in job opportunities (21).

Urbanization can have both positive and negative effects on the health of populations. On average, urban dwellers have better access to education, health care, and favorable living conditions than rural dwellers, and are more likely to be literate and have longer lifespans (19). Such is the case in India, where increased f urbanization in each state is positively correlated with overall economic development and decreasing rates of poverty over the past several decades (20). The benefits that come with living in cities are not evenly distributed, however.

An analysis of several Indian censuses has found that although urbanization was positively correlated with access to healthcare facilities, clean water, and latrines, the correlation has steadily decreased over the past several decades. Increasing urban populations strain the ability of cities to provide basic amenities, decreasing the quality of life for urban dwellers. (21). This is increasingly common especially in informal settlements, commonly known as slums. Slums develop when the influx of people to an urban center is so great that is outstrips the city's ability to accommodate it, leading people to settle wherever they can find space, often on the city's outskirts or in neglected areas. India is no exception: Dharavi, a slum within the city of Mumbai, houses around 700,000 people and accounts for slightly less than half of the city's total population (22). A former mangrove swamp, it is now the second-largest slum in Asia and the third-largest slum in the world (23). It is also one of the world's densest areas, with a population density of over 277,000 people per square mile, more than ten times the population density of New York City (24). However, India has slums across the country, in population centers both large and small. In 2001, the national census for the first time collected data on slum populations. The 2011 census estimated that over 65 million people in 14 million households, over 16% of India's total urban population (20), live in slums. This represents a substantial increase from 2001, in which 40 million people, or 14% of the urban population, were identified as slum dwellers (25). Analysis has shown that slums are more likely to develop in cities undergoing rapid industrialization with more employment opportunity. The potential for employment draws migrants from surrounding areas, straining the capacities of the urban center and leading to the creation of low resource settlements that develop into slums (21).

As is the case in the rest of the developing world, Indian slums are crowded and unsanitary. Since the vast majority of residences in slums are not legally sanctioned, residents constantly face the threat of eviction, though some slums are large enough to have metered electricity and water (26). The dismal condition of slums points to a newly growing issue, in India and the rest of the world: the intersection of urbanization and disease.

## The Intersection of Urbanization and Disease

There is growing evidence that urbanization can be detrimental to public health, especially in developing countries, though the exact relationship between the urban environment and the health of its inhabitants is still uncertain. Several pathogens that were most commonly found in rural environments have emerged or re-emerged in urban areas (27). Malaria was the first of these transitional pathogens to be identified in Freetown, Sierra Leone, in 1899, but other tropical diseases such as soil-transmitted helminths, dengue, and cholera have all become more common in urban settings (28). This is due to a number of factors that are common to cities in developing countries across the globe, the first of which is crowding. The sheer size and density of many cities and especially slums dramatically increases the contact rate of residents, and so facilitates the spread of pathogens , especially via respiratory or fecal-oral routes (19). Urban populations are also in a state of constant flux, and incoming migrants from rural areas or other cities carry new pathogens or reintroducing those that had previously faded out. Moreover, animals such as pigs, dogs, cows, and chickens are either kept as pets and livestock or roam free in cities in many developing countries, increasing the prevalence of zoonoses (29).

Another important factor contributing to urban disease is sanitation. In past decades, communicable disease was more common in rural areas than urban ones, because healthcare, clean water, and sanitary waste disposal were more available in cities. With the increasing prevalence of cities and particularly slums in developing countries, poor sanitation issues are now as common or more common in cities as outside them. A lack of sanitary waste disposal in particular contributes to disease, especially diarrheal disease. In rural areas, both human and animal waste is left to decompose in fields or is burned as fuel (27). These options are not available in urban environments, and so waste is deposited in empty lots or gutters where it encourages the breeding of insects and other potential disease vectors. This waste may also contaminate the water residents use for drinking, further spreading disease (27).

Another factor contributing to the spread of disease in cities is serious and widespread social disparity. Healthcare resources tend to be concentrated in the areas with the most wealth. Impoverished areas, which are already more prone to disease due to crowding and poor sanitation, have the additional burden of being less likely to benefit from city resources (27). A survey of Chandigarh, India, compared healthcare indicators such as vaccination rates and reproductive and child health services among urban, slum, and rural areas, and found that all indicators were at their lowest levels in slum areas, compared to both urban and rural areas (30). Additionally, a survey of urban Bangladesh found that slum children are more likely to suffer from diarrhea, dehydration, and malnutrition than children from other areas (14).

In theory, urbanization could contribute to an overall increase in public health, but the proliferation of slums and their associated hazards may outweigh any potential gains. That urbanization may represent an overall detriment to public health is well documented (12, 14, 27, 30), but the extent of this detriment is not known, for several possible reasons.

There are few studies of health in India that could be used to quantify the effect of urbanization over time. India's enormous population and lack of infrastructure make it difficult to reach some populations and gather a nationally representative sample. Additionally, though there are many cross-sectional or case-control studies, there are few longitudinal studies conducted at the national level. The interpretation of urbanization as a public health issue is also relatively new; though India has been undergoing rapid urbanization since its independence in 1947 (21, 31), the increasing pace of urbanization in developing countries worldwide has become a major issue relatively recently

The aim of this paper is to examine the prevalence of diarrhea in urban and rural centers in India over time to determine if it is meaningfully correlated with urbanization, and if so, what other variables are significantly associated with diarrheal disease in each setting.  Better understanding of the causes of this major health burden can aid in the development of prevention measures.

# Chapter II

## Abstract

India is home to almost one-fifth of the world's population and a third of the world's poor. There is a persistent health gap between rural populations, which typically have worse health outcomes, and rapidly growing urban centers, though the growth of urban slums has complicated the picture. Diarrheal illness is one of the world's leading cause of death and disability and is a major health concern in India. This paper uses data from two waves of the India Human Development Survey, IHDS-I and IHDS-II, to examine prevalence of diarrhea in urban and rural areas of India, as well as associated risk factors. Logistic regression is used for each separate survey wave, and generalized estimating equations are used to conduct longitudinal analysis on data from both survey waves combined. Univariate analysis showed urban environment to have a protective effect on diarrhea in IHDS-I (OR = 0.63, 95% CI: 0.59, 0.67), IHDS-II (OR = 0.78, 95% CI: 0.72, 0.84), and longitudinal analysis (OR = 0.71, 95% CI: 0.67, 0.75). However, in all three cases, the association became null when adjusted for household income, literacy, and piped water, suggesting that these are major contributing factors in the difference between the health of urban and rural populations in India. This paper uses longitudinal data to pinpoint important factors in the public health gap between urban and rural areas in a developing nation and adds to a growing body of literature implicating income inequality as having a negative effect on the health of nations, especially developing nations.

## Introduction

India has one-fifth of the global population, one-third of the world's poor, and one of the world's fastest-growing economies (1, 2). India also has higher disease burden than other countries with similar economic indicators (3). In 2016, communicable diseases were three of the top five contributors to mortality in India. Diarrheal disease is the most common communicable contributor to death and disability, and the third most common contributor overall (4). Diarrheal pathogens are most commonly spread via the fecal-oral route (9-11), and so most risk factors stem from unclean water and unsafe waste disposal in conjunction with person-to-person transmission (12, 32, 33).

India is working to address the issue of diarrhea by improving access to safe water and improving childhood nutrition (34). The Global Burden of Disease project estimates that, between 2005 and 2015, Indian deaths from diarrhea have decreased by 60% overall, and by 88% in children under five (3), though the decline in incidence has not been nearly as steep (8). India's disease burden from non-communicable diseases overtook the burden from communicable, nutritional, maternal, and neonatal diseases for the first time in 2003 (3). The reduction in morbidity from communicable diseases, and diarrhea in particular, can partially be attributed to a focus on providing improved latrines and clean drinking water in impoverished areas (15).

India's transitionary state is also reflected is the process of urbanization. The UN reported in 2014 that, for the first time, over half of the world's population was living in cities (19). As populations increase and the global economy shifts away from agriculture, the proportion of people living in cities will increase, and most growth will occur in developing countries. In 2001, 28% of India's population lived in cities (25); that figure grew to 31% by 2011 (20). By 2050, half of India is projected to live in urban environments (21).

Urbanization can have both positive and negative effects on population health. On average, urban dwellers have better access to education, health care, and favorable living conditions than rural dwellers, and are more likely to be literate and have longer lifespans (19). Such is the case in India, where the percentage of urbanization in each state is positively correlated with the overall level of economic development, and the percentage of urban dwellers below the poverty line has decreased over the past several decades (21). The benefits that come with living in cities are not evenly distributed, however.

As cities grow, so to do slums, especially in developing nations. Slums develop when the influx of people to an urban center is so great that the city cannot grow apace, and people settle on the city's outskirts or in neglected urban areas. In 2001, the Indian national census estimated that 65 million people, 16% or India's urban population, lived in slums (20).

The relationship between communicable disease and the urban environment is still evolving. Several pathogens that were most commonly found in rural environments have emerged or re-emerged in urban areas (27), such as malaria, soil-transmitted helminths, dengue, and cholera (28). The spread of these diseases is due to factors common to cities in developing countries across the globe, the first is of which is crowding. The size and density of many cities and slums means that residents have a high contact rate, making the spread of pathogens via respiratory or fecal-oral routes easier (19). Animals such as pigs, dogs, cows, and chickens are kept as pets and livestock or roam free in cities in many developing countries, increasing the prevalence of zoonoses (29). A lack of sanitary waste disposal also contributes to disease, especially diarrheal disease. In rural areas, human and animal waste decomposes in fields or is burned as fuel (27). These options are not available in cities, so waste is deposited in empty lots or gutters if the sewage system is inadequate or absent. There, it encourages the breeding of insects and other vectors and may contaminate drinking water. Social disparity in cities concentrates healthcare resources in the areas with the most wealth. Impoverished areas, which are already more prone to disease, have the additional burden of being less likely to receive treatment (30).

Urbanization could lead to a decrease in the risk of both communicable and non-communicable disease for urban dwellers, but the proliferation of slums and their associated hazards may result in a detriment to public health. This paper aims to examine the prevalence of diarrhea in urban and rural centers in India over time to determine if it is meaningfully correlated with urbanization, and if so, which variables are significantly associated with diarrheal disease in each setting.

## Methods

*Population sampling and questionnaire*

Data were drawn from the India Human Development Survey (IHDS), a nationally representative survey covering all Indian states and union territories except for Andaman, Nicobar, and Lakshadweep. Households across 1503 rural villages and 971 urban blocks were selected for inclusion using stratified random sampling. Participating households took part in two interviews. The first was conducted with a knowledgeable informant, usually the head of the household, regarding socioeconomic factors such as income, employment, and social capitol. A second interview regarding family topics such household health, education, and family planning, was conducted with an ever-married female between the ages of 15 and 49. The survey was conducted in two waves. The first (IHDS-I) and second waves (IHDS-II) were conducted between November 2004 to October 2005 and November 2011 to October 2012, respectively. Whenever possible, households from IHDS-I were re-contacted and re-interviewed in IHDS-II. If members of an IHDS-I household had split into multiple households, all split household still residing in the same sampling unit (village or urban block) were interviewed. If households were lost due to migration, attrition, or, death, additional households were randomly selected from the same sampling unit as replacements. In all, 85% of households and 70% of individuals from IHDS-I were re-interviewed in IHDS-II. Most information in the survey was self-reported by the interviewees, though some questions relied on observations from the interviewer. IHDS-I and IHDS-II data are publicly available for download through the Inter-university Consortium for Political and Social Research (ICPSR, available at https://www.icpsr.umich.edu/icpsrweb/DSDR/series/507).

The primary outcome in this analysis was an interviewee reporting that a specific household member experienced diarrhea within 30 days prior to the interview. The primary exposure was living in an urban environment. In each survey wave, households were categorized as urban or rural based on the most recent census at the time of the survey; this information was included in the dataset. The census classified an area as urban if it had a municipality, corporation, cantonment board, or town area committee, or if it met the following criteria: a minimum population of five thousand, at least 75% of male main working population engaged in non-agricultural pursuits, and a population density of at least 400 persons per square kilometer (20).

*Data Analysis*

Potential predictors and confounders related to water, sanitation, and hygiene were chosen according to biological plausibility and after a literature review. Factors analyzed included age, highest household education, household income, living below the poverty line, literacy, presence of fever within the last 30 days, presence of cough within the last 30 days, storing water inside vs outside the house, presence or absence of a household toilet, presence or absence of piped indoor drinking water, and practice of water purification. The variables concerning household toilets and water purification were originally reported with several possible responses and were condensed to binomial variables for purposes of analysis. For example, when asked if they purified water, respondents could choose "never," "rarely," "usually," or "always." These responses were then recorded as "no" if the respondent answered "never," and as "yes" if the respondent answered otherwise. The variables concerning household toilets allowed respondents to differentiate which type of toilet they have access to; the binomial variable simplifies responses into "access to toilet" and "no access to toilet." A variable was included that indicated which survey wave a particular observation was from, and another allowed for interaction between the primary outcome and time. Models containing subsets of the above covariates were compared to a model containing all covariates. The final model chosen was the most parsimonious one that produced a measure of association that was within 10% of the measure of association produced by the fully adjusted model in all subsets of analysis. Once a model was identified, a student's t-test was conducted on all predictor variables to determine if prevalence varied significantly between urban and rural areas in both IHDS-I and IHDS-II.

Regression analysis was conducted using data from all participants from each survey wave, and with the subset of data only from participants who took part in both waves. Logistic regression was used for analysis of individual waves. Generalized estimating equations (GEEs) were used for longitudinal analysis. Correlation arising from repeated measures was accounted for using by specifying an unstructured covariance matrix. Because there were only two measurements per person, a more simplified matrix structure would not have produced different results. A logit link function was used to allow for logistic regression on correlated data. Results were expressed as odds ratios (ORs) with corresponding 95% confidence intervals (95% CIs).

A sensitivity analysis was conducted to determine if any adjustment was necessary to correct for recall bias. Validation studies of recall-based diarrhea prevalence measures suggest that recall accuracy drops after just two days and drops considerably after one week (35, 36). They also suggest that recall is unlikely to drop after two weeks; that is, recall stabilizes. As recall fades over time, diarrhea is increasingly under-reported, resulting in false negatives and thus reducing sensitivity. Over-reporting, resulting in false positives, is rare, if it occurs at all, and so specificity remains high. Recall bias is expected to be non-differential between urban and rural respondents and between those who did and did not experience diarrhea. A multi-dimensional bias analysis was conducted, in which odds ratios were examined under sensitivity measures ranging from 0.65 to 0.50, and specificity ranging from 1.0 to 0.98. Because specificity is the primary driver of bias, adjusted measures are identical or very similar to the crude measure of association, even at the lowest measures of sensitivity.

## Results

Table 1: Characteristics of participants in IHDS-I and IHDS-II

| | IHDS-I | | |
|---|---|---|---|
| Characteristics | Rural *n*=143,374 (%) | Urban *n*=72,380 (%) | Total *n*=215,754 (%) |
| Age (years) | | | |
| 0 to 4 | 13,654 (9.5) | 5,834 (8.1) | 19,488 (9.0) |
| 5 to 9 | 16,447 (11.5) | 7,125 (9.8) | 23,572 (10.9) |
| 10 to 14 | 17,559 (12.3) | 7,843 (10.8) | 25,402 (11.8) |
| 15 to 19 | 14,876 (10.4) | 7,717 (10.7) | 22,593 (10.5) |
| 20 to 29 | 23,719 (16.5) | 13,483 (18.6) | 37,202 (17.2) |
| 30 to 39 | 19,254 (13.4) | 10,974 (15.2) | 30,228 (14.0) |
| 40 to 59 | 25,218 (17.6) | 14,147 (19.6) | 39,365 (18.3) |
| 60+ | 12,647 (8.8) | 5,257 (7.3) | 17,904 (8.3) |
| | | | |
| Highest household education (standard grades) | | | |
| None | 33,748 (23.6) | 6,659 (9.2) | 40,407 (18.8) |
| 1 to 4 | 12,065 (8.4) | 3,149 (4.4) | 15,214 (7.1) |
| 5 to 9 | 48,843 (34.1) | 20,068 (27.8) | 68,911 (32.0) |
| 10 to 11 | 20,890 (14.6) | 12,349 (17.1) | 33,239 (15.4) |
| 12 to some college | 14,155 (9.9) | 9,917 (13.7) | 24,072 (11.2) |
| College graduate | 13,486 (9.4) | 20,179 (27.9) | 33,665 (15.6) |
| | | | |
| Annual Household income (rupees) | | | |
| 1st quintile | 37,751 (26.3) | 5,706 (7.9) | 43,457 (20.1) |
| 2nd quintile | 33,631 (23.5) | 9,266 (12.8) | 42,897 (19.9) |
| 3rd quintile | 28,425 (19.8) | 14,677 (20.3) | 43,102 (19.9) |
| 4th quintile | 23,431 (16.3) | 19,713 (27.2) | 43,144 (20.0) |
| 5th quintile | 20,136 (14.0) | 23,018 (31.8) | 43,154 (20.0) |
| | | | |
| Living under poverty line | 34,557 (24.1) | 15,929 (22.0) | 50,486 (23.4) |
| Literate | 82,723 (58.0) | 54,341 (75.4) | 137,064 (63.8) |
| Piped water present in home | 43,379 (30.3) | 51,313 (71.0) | 94,692 (43.9) |
| Female sex | 70,454 (49.1) | 35,495 (49.0) | 105,949 (49.1) |
| Diarrhea reported in past month | 3,903 (2.7) | 1,259 (1.7) | 5,162 (2.4) |

| | IHDS-II | | |
|---|---|---|---|
| Characteristics | Rural<br>*n*=135,116 (%) | Urban<br>*n*=64,611 (%) | Total<br>*n*=204,565 (%) |
| Age (years) | | | |
| 0 to 4 | 12,187 (9.0) | 4,838 (6.0) | 17,025 (8.3) |
| 5 to 9 | 13,333 (9.9) | 5535 (8.0) | 18,868 (9.2) |
| 10 to 14 | 14,807 (11.0) | 6,664 (9.6) | 21,471 (10.5) |
| 15 to 19 | 12,810 (9.5) | 6,465 (9.3) | 19,275 (9.4) |
| 20 to 29 | 22,945 (17.0) | 13,005 (18.7) | 35,950 (17.6) |
| 30 to 39 | 17,734 (13.1) | 10,201 (14.7) | 27,935 (13.7) |
| 40 to 59 | 26,420 (19.6) | 15,699 (22.6) | 42,119 (20.6) |
| 60+ | 14,880 (11.0) | 7,042 (10.1) | 21,922 (10.7) |
| | | | |
| Highest household education (standard grades) | | | |
| None | 24,882 (18.4) | 4,376 (6.3) | 29,258 (14.3) |
| 1 to 4 | 8,679 (6.4) | 2,352 (3.4) | 11,031 (5.4) |
| 5 to 9 | 47,566 (35.2) | 18,276 (26.3) | 65,842 (32.2) |
| 10 to 11 | 20,096 (14.9) | 10,849 (15.6) | 30,945 (15.1) |
| 12 to some college | 18,074 (13.4) | 11,791 (17.0) | 29,865 (14.6) |
| College graduate | 15,802 (11.7) | 21,794 (31.4) | 37,596 (18.4) |
| | | | |
| Annual household income (Rupees) | | | |
| 1st quintile | 10,645 (7.9) | 1,548 (2.2) | 12,193 (6.0) |
| 2nd quintile | 11,837 (8.8) | 1,482 (2.1) | 13,319 (6.5) |
| 3rd quintile | 22,821 (16.9) | 5,083 (7.3) | 27,904 (13.6) |
| 4th quintile | 35,821 (26.5) | 14,717 (21.2) | 50,538 (24.7) |
| 5th quintile | 53,995 (40.0) | 46,620 (67.1) | 100615 (49.2) |
| | | | |
| Living under poverty line | 31,148 (23.1) | 8,748 (12.6) | 39,896 (19.5) |
| Literate | 84,935 (62.9) | 54,915 (79.1) | 139,850 (68.4) |
| Piped water present in home | 28,040 (20.8) | 35,299 (51.0) | 63,339 (31.1) |
| Female sex | 67,940 (50.3) | 34,566 (49.8) | 102,506 (50.1) |
| Diarrhea reported in past month | 3,912 (2.9) | 1,531 (2.2) | 5,443 (2.7) |

Descriptive results

The study population consisted of participants in the IHDS-I collected information from 215,754 individuals in 41,544 households. IHDS-II collected information 204,565 individuals in 42,152 households. 150,988 individuals (70% of IHDS-I) in 40,018 households participated in both rounds. 64,766 individuals (30%) in 6,911 households were lost to attrition, death, or migration and so were only interviewed in IHDS-I. 53,580 individuals (26% of IHDS-II) in 2,134 households were only surveyed in IHDS-II, as replacements for those lost in IHDS-I or as new members of newly selected households. Nineteen primary sampling units (villages/urban blocks) changed from rural to urban between IHDS-I and IHDS-II; no areas changed from urban to rural. In IHDS-I, 14,820 households (36%) were from urban areas; 14,573 (35%) of IHDS-II households were urban.

A diarrhea incidence of 2.39% (5,162/215,754) was reported in IHDS-I, with a lower prevalence in urban (incidence = 1.74%, 1,259/72,380) than rural individuals (2.71%, 3,903/143,374). An incidence of 2.66% (5,443/204,565) was reported in IHDS-II, with a lower prevalence in urban (incidence = 2.20%, 1,531/64,611) than rural individuals (2.90%, 3,912/135,116).

A sensitivity analysis determined that the instrument used to measure diarrhea recall had a high sensitivity (0.98 – 1.00) and a moderate specificity (0.500 – 0.65). Because sensitivity is the primary driver of bias, the measures collected in the original surveys do not require adjustment to account for recall bias.

Table 2: Prevalence of diarrhea by strata of covariates in IHDS-I and IHDS-II

| | IHDS-I | | |
|---|---|---|---|
| Covariates | Rural n=143,374 (%) | Urban n=72,380 (%) | Total *n*=215,754 (%) |
| Age (years) | | | |
| 0 to 4 | 1,089 (9.5) | 368 (7.5) | 1,457 (8.88) |
| 5 to 9 | 511 (3.1) | 151 (2.1) | 662 (2.81) |
| 10 to 14 | 352 (2.0) | 95 (1.2) | 447 (1.76) |
| 15 to 19 | 214 (1.4) | 86 (1.1) | 300 (1.33) |
| 20 to 29 | 311 (1.3) | 121 (0.9) | 432 (1.16) |
| 30 to 39 | 326 (1.7) | 99 (0.9) | 425 (1.41) |
| 40 to 59 | 539 (2.1) | 174 (1.2) | 713 (1.81) |
| 60+ | 308 (2.4) | 68 (1.3) | 376 (1.21) |
| Highest household education (standard grades) | | | |
| None | 1,150 (3.4) | 155 (2.3) | 1,305 (3.23) |
| 1 to 4 | 388 (3.2) | 72 (2.3) | 460 (3.02) |
| 5 to 9 | 1,330 (2.7) | 386 (1.9) | 1,716 (2.49) |
| 10 to 11 | 482 (2.3) | 209 (1.7) | 691 (2.08) |
| 12 to some college | 284 (2.0) | 136 (1.4) | 420 (1.74) |
| College graduate | 266 (2.0) | 300 (1.5) | 566 (1.68) |
| Annual household income (Rupees) | | | |
| 1st quintile | 1,351 (3.6) | 139 (2.4) | 1490 (3.43) |
| 2nd quintile | 1,041 (3.1) | 299 (2.5) | 1,270 (2.96) |
| 3rd quintile | 695 (2.5) | 278 (1.9) | 973 (2.26) |
| 4th quintile | 474 (2.0) | 308 (1.6) | 782 (1.81) |
| 5th quintile | 342 (1.7) | 305 (1.3) | 647 (1.50) |
| Living under poverty line | 3,008 (2.8) | 961 (1.7) | 3,969 (2.40) |
| Not living under poverty line | 895 (2.6) | 298 (1.9) | 1,193 (2.36) |
| Illiterate | 2,452 (4.1) | 646 (3.6) | 3,098 (3.99) |
| Literate | 1,442 (1.7) | 611 (1.1) | 2,053 (1.50) |
| Piped water absent | 3,249 (3.3) | 498 (2.4) | 3,747 (3.10) |
| Piped water present | 654 (1.5) | 761 (1.5) | 1,415 (1.49) |
| Male Sex | 1,944 (2.7) | 600 (1.6) | 2,544 (2.32) |
| Female Sex | 1,959 (2.8) | 659 (1.9) | 2,618 (2.47) |

| | IHDS-II | | |
|---|---|---|---|
| Covariates | Rural *n=135,116 (%)* | Urban n=64,611 (%) | Total *n=204,565 (%)* |
| Age (years) | | | |
| 0 to 4 | 913 (9.1) | 321 (8.0) | 1,234 (8.77) |
| 5 to 9 | 342 (2.6) | 173 (3.1) | 515 (2.73) |
| 10 to 14 | 278 (1.9) | 120 (1.8) | 398 (1.85) |
| 15 to 19 | 233 (1.8) | 88 (1.4) | 321 (1.67) |
| 20 to 29 | 331 (1.4) | 180 (1.4) | 511 (1.42) |
| 30 to 39 | 328 (1.9) | 131 (1.3) | 459 (1.64) |
| 40 to 59 | 687 (2.6) | 275 (1.8) | 962 (2.28) |
| 60+ | 511 (3.4) | 144 (2.0) | 655 (2.99) |
| | | | |
| Highest household education- standard grades | | | |
| None | 841 (3.4) | 135 (3.1) | 976 (3.34) |
| 1 to 4 | 263 (3.0) | 66 (2.8) | 329 (2.98) |
| 5 to 9 | 1,455 (3.1) | 507 (2.8) | 1,962 (2.98) |
| 10 to 11 | 535 (2.7) | 252 (2.3) | 787 (2.54) |
| 12 to some college | 472 (2.6) | 243 (2.2) | 715 (2.39) |
| College graduate | 345 (2.2) | 328 (1.5) | 673 (1.79) |
| | | | |
| Annual household income (Rupees) | | | |
| 1st quintile | 407 (3.8) | 35 (2.3) | 442 (3.63) |
| 2nd quintile | 472 (4.0) | 49 (3.3) | 521 (3.91) |
| 3rd quintile | 800 (3.5) | 155 (3.1) | 955 (3.42) |
| 4th quintile | 1,029 (2.9) | 406 (2.8) | 1,435 (2.84) |
| 5th quintile | 1,204 (2.2) | 886 (1.9) | 2,090 (2.08) |
| | | | |
| Living under poverty line | 3,116 (3.0) | 1,322 (2.2) | 4,438 (2.70) |
| Not living under poverty line | 794 (2.6) | 209 (2.4) | 1,003 (2.51) |
| | | | |
| Illiterate | 2,291 (4.6) | 639 (4.4) | 2,930 (4.54) |
| Literate | 1,619 (1.9) | 892 (1.6) | 2,511 (1.80) |
| | | | |
| Piped water absent | 3,240 (3.0) | 822 (2.4) | 4,062 (2.89) |
| Piped water present | 668 (2.4) | 708 (2.0) | 1,376 (2.17) |
| | | | |
| Male Sex | 1,845 (2.8) | 718 (2.1) | 2,563 (2.51) |
| Female Sex | 2,067 (3.0) | 813 (2.4) | 2,880 (2.81) |

**Table 3:** Difference of covariates between survey waves and results of student's t-test

| | IHDS-I | | | | |
|---|---|---|---|---|---|
| | Mean/prevalence (SD) | | Difference (SD) | t-score | p |
| Covariates | Rural | Urban | | | |
| Piped water in home | 0.30 (0.46) | 0.71 (0.45) | 0.41 (0.46) | -194.95 | <0.0001 |
| Literacy | 0.59 (0.49) | 0.75 (0.43) | 0.17 (0.47) | -80.41 | <0.0001 |
| Diarrhea | 0.027 (0.16) | 0.017 (0.13) | 0.010 (0.15) | 14.11 | <0.0001 |
| | **IHDS-II** | | | | |
| | Mean/prevalence (SD) | | Difference (SD) | t-score | p |
| Covariates | Rural | Urban | | | |
| Piped water in home | 0.21 (0.41) | 0.51 (0.50) | 0.30 (0.44) | -146.83 | <0.0001 |
| Literacy | 0.63 (0.48) | 0.79 (0.41) | 0.16 (0.46) | -75.67 | <0.0001 |
| Diarrhea | 0.029 (0.17) | 0.022 (0.15) | 0.007 (0.16) | 9.19 | <0.0001 |

**Table 4**: Results of unadjusted and adjusted analysis for IHDS-I, IHDS-II, and longitudinal data

| | IHDS-I | |
|---|---|---|
| | Unadjusted | Adjusted |
| | OR (95% CI) | OR (95% CI) |
| Urban residence | 0.63 (0.59, 0.67) | 1.05 (0.98, 1.13) |
| Annual household income | | |
| 1st quintile | 1.00 | 1.00 |
| 2nd quintile | 1.33 (0.24, 0.41) | 0.90 (0.83, 0.97) |
| 3rd quintile | 0.93 (0.87, 1.00) | 0.75 (0.69, 0.82) |
| 4th quintile | 0.71 (0.66, 0.77) | 0.69 (0.63, 0.76) |
| 5th quintile | 0.57 (0.52, 0.62) | 0.64 (0.58, 0.70) |
| Literate | 0.37 (0.35, 0.39 | 0.42 (0.39, 0.44) |
| Piped water present | 0.47 (0.47, 0.50) | 0.56 (0.52, 0.60) |
| | **IHDS-II** | |
| | Unadjusted | Adjusted |
| | OR (95% CI) | OR (95% CI) |
| Urban residence | 0.76 (0.71, 0.80) | 1.01 (0.95, 1.08) |
| Income- 1st quintile | 1.00 | 1.00 |
| Income- 2nd quintile | 1.54 (1.41, 1.69) | 1.08 (0.95, 1.23) |

| | | |
|---|---|---|
| Income- 3rd quintile | 1.36 (1.27, 1.46) | 0.97 (0.86, 1.08) |
| Income- 4th quintile | 1.09 (1.03, 1.16) | 0.84 (0.75, 0.93) |
| Income- 5th quintile | 0.64 (0.60, 0.67) | 0.68 (0.61, 0.76) |
| Literate | 0.38 (0.36, 0.41) | 0.41 (0.39, 0.43) |
| Piped water present | 0.75 (0.70, 0.79) | 0.93 (0.87, 0.99) |
| | Longitudinal Analysis | |
| | Unadjusted | Adjusted |
| | OR (95% CI) | OR (95% CI) |
| Urban residence | 0.71 (0.67, 0.75) | 0.90 (0.76, 1.07) |
| Income- 1st quintile | 1.00 | 1.00 |
| Income- 2nd quintile | 1.39 (1.31, 1.48) | 0.89 (0.82, 0.96) |
| Income- 3rd quintile | 1.08 (1.02, 1.15) | 0.76 (0.71, 0.83) |
| Income- 4th quintile | 0.89 (0.84, 0.94) | 0.70 (0.65, 0.76) |
| Income- 5th quintile | 0.60 (0.55, 0.62) | 0.58 (0.54, 0.63) |
| Literate | 0.45 (0.43, 0.47) | 0.51 (0.49, 0.54) |
| Piped water present | 0.60 (0.57, 0.63) | 0.68 (0.64, 0.72) |
| Response recorded in IHDS-I | 1.00 | 1.00 |
| Response recorded in IHDS-II | 0.79 (0.76, 0.83) | 0.97 (0.92, 1.02) |

Analysis results

Living in an urban environment had a significant protective effect on diarrhea in univariate analysis in IHDS-I (OR = 0.63, 95% CI: 0.59, 0.67), in IHDS-II (OR = 0.76, 95% CI: 0.71, 0.80), and in longitudinal analysis (OR = 0.71, 95% CI = 0.67, 0.75). However, in all cases, the association was nullified when adjusting for annual household income, literacy, and piped water (Table 4). The association of diarrhea with covariates was also consistent across all IHDS-I, IHDS-II, and longitudinal analysis. Literacy and piped water were both significantly protective in both unadjusted and adjusted analysis. Those in lower income quintiles had higher odds of diarrhea. As income increased, odds of diarrhea decreased correspondingly. Univariate and multivariate analysis of all variables considered is included in the appendix (Table S1). We found no evidence of interaction between urban environment and time.

## Discussion

In IHDS-I, IHDS-II, and longitudinal analysis, the urban environment was strongly associated with decreased odds of diarrhea. However, this effect became very close to null when accounting for the effects of income, literacy, and piped water, all factors that have been shown to have a significant impact on health, and diarrhea in particular, in similar settings (14, 30, 37, 38).

Literacy and income might both be considered proxy variables for socioeconomic status and geographic area of residence (37). Education has been strongly linked to health outcomes in developed and developing countries. While the highest level of household education and years of individual education were both considered for analysis in this study, they did not end up being as significant as individual literacy, which in this case may function as a dichotomous absent/present education variable and may represent the threshold at which the protective effects of education take effect. Income has a similarly well-established relationship with health. Lower income has direct physiological effects that stem from low quality housing, poor nutrition, increased air pollution, and the psychological stress of poverty, to name just a few examples. Income is also inextricably linked with geographic area of residence. As noted in the introduction to this paper, slums house the poorest of each city's residents. They therefore have the worst conditions, which in turn confer negative health effects. Rural areas tend to be poorer than urban areas in both developing and developed nations. Such was the case in this study: in both survey waves, the mean income of urban areas was substantially higher than that of rural areas.

Water quality is an issue of much concern in rural India. Rural dwellers, lacking improved water sources, make use of unimproved surface sources such as lakes and rivers (39). Such sources are vulnerable to contamination from unsanitary waste disposal practices, including open defecation, which has been strongly linked to diarrhea disease (40, 41). Open defecation is vastly more common in rural areas: in the 2011 census, 67% of respondents reported practicing open defecation, while the corresponding number in urban areas was just 13% (20). Piped water, in contrast to surface water, is more resistant to contamination. Enclosed pipes protect from surface runoff, and since the system is typically under continuous pressure, any breaches lead to water flowing out but not pollutants flowing in. The water may still be subject to contamination at its source, but it is often the best option available and has been linked to positive health outcomes in several studies (37). In both survey waves in this study, piped water was more than twice as common in urban areas. India's government recognizes this issue and has been investing in providing water services to rural communities since the 1950's (42). More recently, in 2009, the National Rural Drinking Water Programme was launched to renovate current piped water facilities, many of which are in disrepair, and further invest in the creation of new ones (43).

The stark contrast between the prevalence of health-promoting factors in urban and rural India may be indicative of the disparity between India's socioeconomic classes. The materialist mechanisms of income's effect on health are obvious, as previously stated, but there is a growing body of literature supporting the notion that relative income matters more than absolute (44-46). Countries with higher wealth disparity tend to have worse overall health outcomes, and India is no exception. Since the 1980s, India's top earners have captured a steadily increasing share of national wealth: in 1982, India's wealthiest 1% accounted for 6.2% of the national GDP, while in 2013 they accounted for 21.7%. India's poorest 50%, in 1982, accounted for 23.6% of the national GDP, but by 2013 that share had fallen to 14.9% (47). India has made progress in many areas, including healthcare and infrastructure, but consistently has worse health outcomes than economically similar countries such as China and Brazil, and has a far greater share of people living in poverty (48).

## Limitations

A limitation of this data is that diarrhea and other health outcomes are consistently under-reported due to falling recall. In addition, most of the measures collected were reported by a single person in the household who reported on the other members, rather than reported by the household members themselves. A sensitivity analysis based on several validation studies suggested that, conditional on the accuracy of the bias model, the ratio measures of association from the crude data can be taken at face value, even though under-reporting is virtually guaranteed, because recall bias is expected to be non-differential between urban and rural, and between those with and without diarrhea methods.

Another limitation of the data is that the subset of longitudinal data containing individuals interviewed in both rounds of data collection is no longer a representative sample of India. The emphasis on households in the follow-up procedure means that individuals who were insolvent or unable to maintain a household were lost to follow-up, while the more stable and socioeconomically well-off individuals remained in the sample. The results should therefore be interpreted with some caution.

## Strengths

The data have several notable strengths, the first of which is the size of the dataset. The IHDS is large, with observations on over two hundred thousand individuals per survey wave. This gives analysis a great deal of power, and so we can be confident that any results are unlikely to be statistical errors or anomalies, conditional on measurement accuracy. The IHDS is also one of few longitudinal studies that take place in India. This allows for a unique insight in a way that cannot be replicated from cross-sectional studies, which are far more common.

## Conclusion

It was conjectured that the urban environment may have an overall harmful effect on public health in the form of increased diarrhea prevalence. That was shown not to be the case; the urban environment has a protective effect on diarrhea incidence, but that effect was nullified when adjusting for piped water, annual household income, and literacy. These, then, are important explanatory factors not only for diarrheal illness but also for broad urban and rural health contexts. Income and education are strongly linked to socioeconomic status, which in turn is strongly linked to health.

Future research may focus on the downstream effects of programs meant to increase the socioeconomic status of individuals or families. Additionally, several of the possible covariates changed in prevalence between IHDS-I and IHDS-II. Piped water, for example, became less common in both urban and rural areas from one survey wave to another. Future studies might examine factors relating to how common certain covariates are, and if they have an impact on health. Studies might also examine this relationship in different states in India, as their levels of economic development vary.

## Chapter III

We found that the urban environment has a protective effect on diarrhea risk, but that this effect is completely explained by higher income, piped water, and greater literacy.  Though measures are almost certainly subject to under-reporting, a bias analysis determined that the calculated odds ratios were valid.  Results were consistent across three separate analyses, lending credibility to findings.

This study adds to the body of literature concerning health disparities between urban and rural populations, especially in developing nations, and identifies specific causal factors for such a disparity.  Literacy and household income, two of the three causal factors, might be considered proxy variables for socioeconomic status, which has repeatedly been shown to have a major effect on health at both the individual and national level.  Piped water, the third causal factor, is a more direct exposure, and has a decidedly beneficial effect on diarrheal disease in this analysis.  The disparity between urban and rural environments in all three causal factors suggests that income inequality plays an important role in the health of nations.

Future research may focus on the downstream effects of programs meant to increase the socioeconomic status of individuals or families.  Additionally, several of the possible covariates changed in prevalence between IHDS-I and IHDS-II. Piped water, for example, became less common in both urban and rural areas from one survey wave to another.  Future studies might examine factors relating to how common certain covariates are, and if they have an impact on health.  Studies might also examine this relationship in different states in India, as their levels of economic development vary.  Additionally, it would be worthwhile to conduct a similar analysis using the Demographic and Health Survey (DHS), another longitudinal nationally representative survey in India that collects information on multiple factors regarding health and socioeconomic status.

# References


1. Fund IM. IMF World Economic Output (WEO) Update, January 2017: A shifting Global Economic Landscape. IMF.org, 2017,
2. Pedro Olinto KB, Carlos Sobrado, Hiroki Uematsu. The state of the poor: where are the poor, where is extreme poverty harder to end, and what is the current profile of the world's poor? *Economic Premise*. Washington, DC: The World Bank, 2013.
3. Dandona L, Dandona R, Kumar GA, et al. Nations within a nation: variations in epidemiological transition across the states of India, 1990–2016 in the Global Burden of Disease Study. *The Lancet*.
4. Evaluation IfHMa. GBD Compare Data Visualization. Seattle, WA: IHME, University of Washington, 2016.
5. Richard SA, Black RE, Gilman RH, et al. Diarrhea in early childhood: short-term association with weight and long-term association with length. *American journal of epidemiology* 2013;178(7):1129-38.
6. Checkley W, Epstein LD, Gilman RH, et al. Effects of acute diarrhea on linear growth in Peruvian children. *American journal of epidemiology* 2003;157(2):166-75.
7. Guerrant DI, Moore SR, Lima AA, et al. Association of early childhood diarrhea and cryptosporidiosis with impaired physical fitness and cognitive function four-seven years later in a poor urban community in northeast Brazil. *The American journal of tropical medicine and hygiene* 1999;61(5):707-13.
8. Estimates of global, regional, and national morbidity, mortality, and aetiologies of diarrhoeal diseases: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Infectious diseases* 2017;17(9):909-48.

9. (CDC) CfDCaP. Rotavirus: Transmission. 201g. (https://www.cdc.gov/rotavirus/about/transmission.html). (Accessed 2017).

10. (CDC) CfDCaP. *Shigella*- Shigellosis. US Department of Health and Human Services; 2017. (https://www.cdc.gov/shigella/general-information.html). (Accessed 2017).

11. (CDC) CfDCaP. *Campylobacter* (Campylobacteriosis). US Department of Health and Human Services; 2017. (https://www.cdc.gov/campylobacter/faq.html). (Accessed 2017).

12. Baker KK, O'Reilly CE, Levine MM, et al. Sanitation and Hygiene-Specific Risk Factors for Moderate-to-Severe Diarrhea in Young Children in the Global Enteric Multicenter Study, 2007-2011: Case-Control Study. *PLoS medicine* 2016;13(5):e1002010.

13. Berendes D, Leon J, Kirby A, et al. Household sanitation is associated with lower risk of bacterial and protozoal enteric infections, but not viral infections and diarrhoea, in a cohort study in a low-income urban neighbourhood in Vellore, India. *Tropical medicine & international health : TM & IH* 2017;22(9):1119-29.

14. Ferdous F, Das SK, Ahmed S, et al. Diarrhoea in slum children: observation from a large diarrhoeal disease hospital in Dhaka, Bangladesh. *Tropical medicine & international health : TM & IH* 2014;19(10):1170-6.

15. Fischer Walker CL, Perin J, Aryee MJ, et al. Diarrhea incidence in low- and middle-income countries in 1990 and 2010: a systematic review. *BMC public health* 2012;12(1):220.

16. Anjana RM, Ali MK, Pradeepa R, et al. The need for obtaining accurate nationwide estimates of diabetes prevalence in India - Rationale for a national study on diabetes. *The Indian Journal of Medical Research* 2011;133(4):369-80.

17. Bureau PI. Government to spend about 2 lakh crore rupees for Scachh Bharat mission more then 11 crore toilets will be built in 5 years., 2014, (Development MoR

18. Tyagi G. Is India Really 'Winning its War' on Open Defecation? [electronic article]. *The Diplomat.* Advance Access: 12/2/2017.

19. WHO UH. *Hidden Cities: Unmasking and overcoming health inequalities in urban settings*. WHO Library; 2010.

20. 2011 Census of India. India, 2011, (Affairs MoH (Office of the Registrar General and Census Commissioner

21. K.C. Sivaramakrishnan AK, B. N. Singh. *Handbook of Urbanization in India*. India: Oxford University Press; 2005.

22. Risbud N. Slums: the case of Mumbai, India [electronic article].

23. Jacobson M. Dharavi, Mumbai's Shadow City [electronic article]. *National Geographic.* Advance Access: 1/10/2018.

24. NYC Population Facts. 2018, (Planning N (City of New York

25. 2001 Census of India. 2001, (Affairs MoH (Office of the Registrar General and Census Commissioner

26. Joanna Lobo VA. Pipe Dreams for Dharavi [electronic article]. *DNA India.* Advance Access: 12/16/2017.

27. Alirol E, Getaz L, Stoll B, et al. Urbanisation and infectious diseases in a globalised world. *The Lancet Infectious diseases* 2011;11(2):131-41.

28. Utzinger J, Keiser J. Urbanization and tropical health--then and now. *Annals of tropical medicine and parasitology* 2006;100(5-6):517-33.

29. Livestock Keeping in Urban Areas: A review of traditional technologies based on literature and field experience [electronic article]. *FAO Animal Production and Health Papers.* Advance Access: 12/2/2017.

30. Gupta M, Thakur JS, Kumar R. Reproductive and child health inequities in Chandigarh Union Territory of India. *Journal of urban health : bulletin of the New York Academy of Medicine* 2008;85(2):291-9.

31. A. Bopegamage MC, V.S. D'Souza et al. *Urbanization in India: Socialogical contribuitons*. Sage Publications; 2003.

32. Santos CA, Strina A, Amorim LD, et al. Individual and contextual determinants of the duration of diarrhoeal episodes in preschool children: a longitudinal study in an urban setting. *Epidemiology and infection* 2012;140(4):689-96.

33. Alemu B, Berhane Y, Worku A. *Prevalence of diarrhea and associated risk factors among children under-five years of age in Eastern Ethiopia: A cross-sectional study*. 2013.

34. Karande S, Gogtay NJ. Impact of the mid-day meal scheme in India. *Journal of postgraduate medicine* 2014;60(2):113-5.

35. Arnold BF, Galiani S, Ram PK, et al. Optimal recall period for caregiver-reported illness in risk factor and intervention studies: a multicountry study. *American journal of epidemiology* 2013;177(4):361-70.

36. Zafar SN, Luby SP, Mendoza C. Recall errors in a weekly survey of diarrhoea in Guatemala: determining the optimal length of recall. *Epidemiology and infection* 2010;138(2):264-9.

37. Solande Desai SA. Maternal education and child health: is there a strong causal relationship? *Demography* 1998;35(1):71-81.

38. Spears D, Ghosh A, Cumming O. Open defecation and childhood stunting in India: an ecological analysis of new data from 112 districts. *PloS one* 2013;8(9):e73784.

39. Ashbolt NJ. Microbial contamination of drinking water and disease outcomes in developing regions. *Toxicology* 2004;198(1-3):229-38.

40. Mara D, Lane J, Scott B, et al. Sanitation and health. *PLoS medicine* 2010;7(11):e1000363.

41. Lanata CF, Huttly SR, Yeager BA. Diarrhea: whose feces matter? Reflections from studies in a Peruvian shanty town. *The Pediatric infectious disease journal* 1998;17(1):7-9.

42. Water Safety Plans for Rural Water Supply in India: Policy Issues and Institutional Arrangements World Bank, 2010.

43. National Rural Drinking Water Programme: Movement towards ensuring people's drinking water security in Rural India. Framework for implementation (updated 2013). India, 2013, (Sanitation MoDWa (Government of India

44. Wilkinson RG. Socioeconomic determinants of health: Health inequalities: relative or absolute material standards? *The BMJ* 1997;314(591).

45. Biggs B, King L, Basu S, et al. Is wealthier always healthier? The impact of national income level, inequality, and poverty on public health in Latin America. *Social science & medicine (1982)* 2010;71(2):266-73.

46. Rajan K, Kennedy J, King L. Is wealthier always healthier in poor countries? The health implications of income, inequality, poverty, and literacy in India. *Social science & medicine (1982)* 2013;88:98-107.

47. Lucas Chancel TP. Indian income inequlaity, 1922-2015: From British Raj to Billionaire raj? In: Database WI, ed, 2017.

48. Horton R, Das P. Indian health: the path from crisis to progress. *Lancet (London, England)* 2011;377(9761):181-3.

APPENDIX A: Supplemental table of odds ratios for all considered covariates

**Table 5:** Unadjusted and adjusted odds ratios for reported diarrhea in IHDS-I and IHDS-II

| | IHDS-I | | | |
|---|---|---|---|---|
| | Unadjusted | | Adjusted | |
| | OR (95% CI) | p | OR, (95% CI) | p |
| Urban Residence | 0.633 (0.59, 0.67) | <.0001 | 1.11 (1.01, 1.216) | 0.0244 |
| Age (Years) | | | | |
| 0 to 4 | 5.88 (5.54, 6.24) | <.0001 | 1 | |
| 5 to 9 | 1.21 (1.11, 1.31) | <.0001 | 0.37 (0.32, 0.42) | <.0001 |
| 10 to 14 | 0.71 (0.64, 0.78) | <.0001 | 0.31 (0.27, 0.36) | <.0001 |
| 15 to 19 | 0.52 (0.46, 0.59) | <.0001 | 0.28 (0.24, 0.33) | <.0001 |
| 20 to 29 | 0.43 (0.39, 0.48) | <.0001 | 0.27 (0.24, 0.31) | <.0001 |
| 30 to 39 | 0.54 (0.49, 0.60) | <.0001 | 0.3 (0.26, 0.34) | <.0001 |
| 40 to 59 | 0.71 (0.66, 0.77) | <.0001 | 0.34 (0.30, 0.38) | <.0001 |
| 60+ | 0.87 (0.78, 0.96) | 0.0076 | 0.35 (0.30, 0.40) | <.0001 |
| Highest household education (standard grades) | | | | |
| none | 1.48 (1.39, 0.58) | <.0001 | 1 | |
| 1 to 4 | 1.30 (1.18, 1.43) | <.0001 | 1.01 (0.89, 1.16) | 0.8316 |
| 5 to 9 | 1.06 (1.00, 1.13) | 0.042 | 0.96 (0.87, 1.06) | 0.4119 |
| 10 to 11 | 0.85 (0.78, 0.92) | <.0001 | 0.84 (0.74, 0.96) | 0.0071 |
| 12 to some college | 0.7 (0.63, 0.77) | <.0001 | 0.75 (0.65, 0.87) | 0.0001 |
| College graduate | 0.66 (0.61, 0.72) | <.0001 | 0.93 (0.81, 1.07) | 0.338 |
| Annual household income (Rupees) | | | | |
| 1st quintile | 1.63 (0.53, 0.73) | <.0001 | 1 | |
| 2nd quintile | 1.325, (0.24, 0.41) | <.0001 | 0.83 (0.76, 0.92) | 0.0001 |
| 3rd quintile | 0.93 (0.87, 1.00) | 0.0402 | 0.74 (0.67, 0.82) | <.0001 |
| 4th quintile | 0.71 (0.66, 0.77) | <.0001 | 0.76 (0.68, 0.86) | <.0001 |
| 5th quintile | 0.57 (0.52, 0.62) | <.0001 | 0.77 (0.67, 0.88) | 0.0001 |
| Living below poverty line | 0.98 (0.92, 1.05) | 0.6215 | 0.86 (0.79, 0.94) | 0.0008 |
| literate | 0.37 (0.35, 0.39) | <.0001 | 0.91 (0.83, 1.00) | 0.0466 |
| Fever in past month | 22.13 (20.84, 23.52) | <.0001 | 8.07 (7.25, 8.98) | <.0001 |
| Cough in past month | 19.25 (18.16, 20.40) | <.0001 | 2.15 (1.94, 2.39) | <.0001 |
| Lidded water storage in home | 1.02 (0.91, 1.15) | 0.7566 | 1.09 (0.95, 1.23) | 0.2136 |
| Toilet present in home | 0.63 (0.60, 0.67) | <.0001 | 0.89 (0.82, 0.97) | 0.0105 |
| Slum | 0.95 (0.86, 1.05) | <.0001 | 1.09 (0.97, 0.21) | 0.1419 |
| Observed excrement outside of household | 1.72 (1.62, 1.82) | <.0001 | 1.26 (0.16, 0.35) | <.0001 |

| | | | | |
|---|---|---|---|---|
| Number of persons in household | 0.99 (0.98, 1.00) | 0.0175 | 1.00 (0.98, 1.01) | 0.6779 |
| Piped water in household | 0.45 (0.45, 0.50) | <.0001 | 0.65 (0.61, 0.71) | <.0001 |
| rarely/sometimes/always purifies water | 0.59 (0.55, 0.63) | <.0001 | 0.91 (0.84, 0.99) | 0.0197 |
| Water stored indoors | 0.92 (0.87, 0.97) | 0.0016 | 1.01 (0.93, 1.09) | 0.8585 |

| | IHDS-II | | | |
|---|---|---|---|---|
| | Unadjusted | | Adjusted | |
| | OR (95% CI) | p | OR (95% CI) | p |
| Urban Residence | 0.76 (0.71, 0.80) | <.0001 | 0.99 (0.91, 1.07) | 0.712 |
| Age (Years) | | | | |
| 0 to 4 | 0.63 (0.59, 0.66) | <.0001 | 1 | |
| 5 to 9 | 1.03 (0.94, 1.13) | 0.5332 | 0.36 (0.31, 0.40) | <.0001 |
| 10 to 14 | 0.67 (0.60, 0.74) | <.0001 | 0.32 (0.28, 0.37) | <.0001 |
| 15 to 19 | 0.60 (0.53, 0.67) | <.0001 | 0.33 (0.28, 0.39) | <.0001 |
| 20 to 29 | 0.48 (0.44, 0.53) | <.0001 | 0.33 (0.29, 0.38) | <.0001 |
| 30 to 39 | 0.56 (0.52, 0.63) | <.0001 | 0.34 (0.30, 0.38) | <.0001 |
| 40 to 59 | 0.82 (0.77, 0.88) | <.0001 | 0.41 (0.37, 0.45) | <.0001 |
| 60+ | 1.14 (1.05, 1.24) | 0.0015 | 0.48 (0.43, 0.53) | <.0001 |
| Highest household education (standard grades) | | | | |
| none | 1.32 (1.23, 1.42) | <.0001 | 1 | |
| 1 to 4 | 1.13 (1.01, 0.27) | 0.0309 | 1.05 (0.91, 1.21) | 0.5203 |
| 5 to 9 | 1.19 (1.13, 1.26) | <.0001 | 1.08 (0.98, 0.19) | 0.1431 |
| 10 to 11 | 0.95 (0.88, 1.02) | 0.1634 | 1.11 (0.99, 1.25) | 0.0873 |
| 12 to some college | 0.88 (0.81, 0.96) | 0.002 | 1.10 (0.97, 1.25) | 0.121 |
| College graduate | 0.62 (0.57, 0.67) | <.0001 | 0.92 (0.81, 1.06) | 0.2368 |
| Annual household income (Rupees) | | | | |
| 1st quintile | 1.41 (1.28, 1.56) | <.0001 | 1 | |
| 2nd quintile | 1.54 (0.41, 0.69) | <.0001 | 0.94 (0.81, 1.09) | 0.3925 |
| 3rd quintile | 1.36, (1.27, 0.46) | <.0001 | 0.90 (0.79, 1.03) | 0.1388 |
| 4th quintile | 1.09 (1.03, 1.16) | 0.004 | 0.88 (0.78, 1.00) | 0.0523 |
| 5th quintile | 0.64 (0.60, 0.67) | <.0001 | 0.81 (0.71, 0.92) | 0.0011 |
| Living below poverty line | 0.93 (0.87, 0.99) | <.0001 | 0.81 (0.75, 0.88) | <.0001 |
| literate | 0.38 (0.36, 0.41) | <.0001 | 0.88 (0.80, 0.95) | 0.0023 |
| Fever in past month | 19.62 (18.41, 20.91) | <.0001 | 14.28 (13.07, 15.60) | <.0001 |
| Cough in past month | 11.12 (10.52, 11.76) | <.0001 | 1.26 (1.17, 1.37) | <.0001 |
| Lidded water storage in home | 1.00 (0.92, 1.09) | 0.9325 | 1.09 (1.00, 0.20) | 0.0552 |
| Toilet present in home | 0.71 (0.67, 0.74) | <.0001 | 0.91 (0.84, 0.98) | 0.0136 |

| | | | | |
|---|---|---|---|---|
| Slum | 0.97 (0.81, 0.18) | <.0001 | 1.19 (0.96, 1.47) | 0.1085 |
| Observed excrement outside household | 1.07 (1.01, 1.14) | 0.0268 | 0.95 (0.88, 1.02) | 0.1394 |
| Number of persons in household | 0.96 (0.95, 0.97) | <.0001 | 1.00 (0.99, 1.02) | 0.5881 |
| Piped water in household | 0.75 (0.70, 0.79) | <.0001 | 1.21 (1.09, 1.34) | 0.0003 |
| rarely/sometimes/always purifies water | 0.71 (0.66, 0.75) | <.0001 | 0.91 (0.85, 0.98) | 0.016 |
| Water stored indoors | 0.76 (0.72, 0.81) | <.0001 | 0.78 (0.71, 0.86) | <.0001 |