Child Mortality from Lower Respiratory Infections and the Effect of Arsenic-contaminated Drinking Water in Rural Bangladesh

by

Warren Christopher Jochem

B.A., University of North Carolina at Chapel Hill, 2007

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Master of Arts

Department of Geography

2012

This thesis entitled:

Child Mortality from Lower Respiratory Infections and the Effect of Arsenic-contaminated Drinking Water in Rural Bangladesh

written by Warren Christopher Jochem

has been approved for the Department of Geography

____

Elisabeth D. Root

____

Fernando Riosmena

____

Randall Kuhn

Date _____________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB protocol # 12-0183

Jochem, Warren Christopher (M.A., Geography)

Child Mortality from Lower Respiratory Infections and the Effect of Arsenic-contaminated Drinking Water in Rural Bangladesh

Thesis directed by Assistant Professor Elisabeth D. Root

Abstract

Children in Bangladesh continue to experience high mortality rates from acute lower respiratory infections (ALRI) despite ongoing treatment programs. Exposure to inorganic arsenic from contaminated drinking water is also a serious health threat in Bangladesh. While lung diseases are commonly associated with inhalation exposures, studies have suggested that exposure to arsenic from drinking water may also negatively affect the lungs. Employing a disease ecology framework, this thesis examines the patterns of mortality from ALRI and the risk factor of arsenic exposure in children under 2 years old between 1989 and 1996 in Matlab, Bangladesh. During this period a community-based treatment program for ALRI was initiated, while arsenic exposure remained high due to widespread well use and the unknown contamination problem. Using a zero-inflated negative binomial regression model, I first examine the association between arsenic exposure from contaminated drinking water and increased mortality rates. Second, I use a spatial scan statistic to test for local clusters of respiratory infection mortality while using the results of the zero-inflated model to adjust the cluster tests. The results suggest that the ALRI treatment program was successful in reducing mortality and that its placement in the region strongly influences the mortality patterns. Arsenic exposure was not significantly associated with ALRI mortality after controlling for the treatment program, socioeconomic status, and access to care. This study focuses on ALRI mortality and cannot rule out a possible association between arsenic and morbidity from respiratory infections. Future work will address this concern as well as the challenges of estimating retrospective exposure to arsenic from drinking water. The continued burden of ALRI makes it important to continue to evaluate intervention programs alongside potential environmental risk factors, such as arsenic exposure, in order to improve child survival.

Acknowledgements

Many thanks go to my advisor, Elisabeth Root, for her continued guidance and support throughout this project and my graduate career. I also thank my committee members, Fernando Riosmena and Randall Kuhn, for their feedback which challenged me to think more deeply about my methodology. Finally, thanks always go to my family for their love and encouragement.

Table of Contents

List of Tables

List of Figures

Introduction

Acute lower respiratory infections (ALRI), such as pneumonia, remain among the leading infectious diseases in the world and disproportionately affect children in many developing countries (Rudan, Boschi-Pinto, Biloglav, Mulholland, & Campbell, 2008). Although these diseases are treatable and preventable, pneumonia is the leading cause of child mortality globally, responsible for 18 percent of all deaths in children under 5 years (R. E. Black et al., 2010). In Bangladesh, ALRI are a leading cause of death among children (Spika et al., 1989) and it is estimated that in 2008 over 25,000 deaths in children under 5 were due to pneumonia (R. E. Black et al., 2010).

ALRI are a class of infections in the lower airways which are caused by both bacterial and viral agents (Simoes et al., 2006). Prior epidemiological studies have found strong relationships between factors such as poverty, malnutrition, and poor living conditions and increased risks of lower respiratory diseases (Lanata & Black, 2008; Rudan et al., 2008). Additionally, some evidence now suggests that exposure to inorganic arsenic in contaminated drinking water may also affect the lungs by decreasing lung function which may lead to chronic lung disease, and an increased susceptibility to respiratory infections (Guha Mazumder, 2007; Raqib et al., 2009). However, the relationship between arsenic and ALRI remains unclear and an associated increased risk of mortality has not been previously studied.

Arsenic is a naturally occurring element in soils which is frequently released into subsurface aquifers. Bangladesh and the wider Bengal Basin Region have a well-documented history of dangerously high levels of arsenic in its drinking water. The widespread exposure is an unintended consequence of a program begun in the 1970s to install shallow tubewells in order to provide clean drinking water and prevent diarrheal disease (Smith, Lingas, & Rahman, 2000).

Arsenic concentrations in the groundwater exhibit high spatial variability due to local geology. At the same, the population's exposure to that arsenic is uneven due to policies and socio-demographic factors that affect access to wells (Caldwell, Caldwell, Mitra, & Smith, 2003). Therefore local-scale spatial variations in health outcomes associated with arsenic are possible due to the variation in exposure from this combination of social and environmental factors; yet these spatial patterns have not been fully explored.

Despite the known risk factors for ALRI, community-based intervention programs designed to provide early case detection and treatment have been shown to be effective at reducing child mortality rates in Bangladesh and in other developing countries (Sazawal & Black, 2003). However, previous studies evaluating the benefits of such treatment programs have not considered the effects of arsenic exposure on ALRI mortality. The continued global burden of ALRI makes it important to evaluate intervention programs alongside potential environmental risk factors, such as arsenic exposure, in order to improve child survival.

Research Objectives

Understanding the relationship between arsenic exposure and respiratory infections demands an understanding of human-environment interactions. With few exceptions (Ali, Emch, Tofail, & Baqui, 2001; Paul, 2004; Paul & De, 2000), geographers have been largely absent from the scientific debate on arsenic and health, and as a result most past studies have not used theories from health and medical geography or more advanced spatial analysis methods. The objective of this research is to explore the geographic patterns of child mortality from lower respiratory infections in the context of an ALRI control program in rural Bangladesh and specifically to answer the questions:

1. Is exposure to arsenic in drinking water associated with higher child mortality from acute lower respiratory infections?
2. Are there significant clusters of child mortality from respiratory infections over space in Matlab, and, if so, where?
3. Are spatial clusters related to arsenic exposure from drinking water, or are they explained by other social and demographic factors?

I hypothesize that exposure to higher levels of arsenic from contaminated wells will be associated with an increase in child mortality from respiratory infections after controlling for the treatment program and other socio-demographic factors. I also hypothesize that the spatial variation in arsenic levels within the study area will at least partially explain the local, spatial patterns of this disease.

As the spatial patterns of child mortality and respiratory infections have received relatively little attention, this research can contribute to the literature in the fields of medical geography and public health. By providing a detailed and spatial study of respiratory infection mortality and arsenic exposure in the context of an ALRI control program this research can also contribute to the design and evaluation of public health intervention strategies. Though this research focuses on Bangladesh, it has broader application in many areas of the world including Chile, Taiwan, and parts of the U.S. which have reported arsenic contaminated groundwater (Smedley & Kinniburgh, 2002). In addition to providing a more complete understanding of the health effects of arsenic exposure, methodologically this work provides an example for employing spatial analysis techniques in future population health studies.

Background

To study the relationship between arsenic and lung disease I focus on the region of Matlab, Bangladesh. Although other areas of Bangladesh suffer from similar health problems, Matlab is a site of ongoing demographic and health surveillance and field studies meaning that specialized data are available for this area, enabling these analyses. As with much of the developing world, this region continues to be burdened by acute lower respiratory infections. This region of Bangladesh also experiences some of the highest rates of arsenic contamination in its drinking water in all of Bangladesh and the world (BGS & DPHE, 2001), making it an important case study to explore this possible relationship.

Study Site: Matlab, Bangladesh

Matlab, a rural region of over 200,000 people is located in central Bangladesh approximately 50 km southeast of the capital city Dhaka (Figure 1). The population is predominantly Muslim with a minority Hindu population. Main occupations include small-scale agriculture and fishing. The household economic and social life is primarily centered around the *bari* – groups of patrilineally-related housing units which share a common courtyard, and, often, a drinking water well. Baris are the primary unit of analysis in this study.

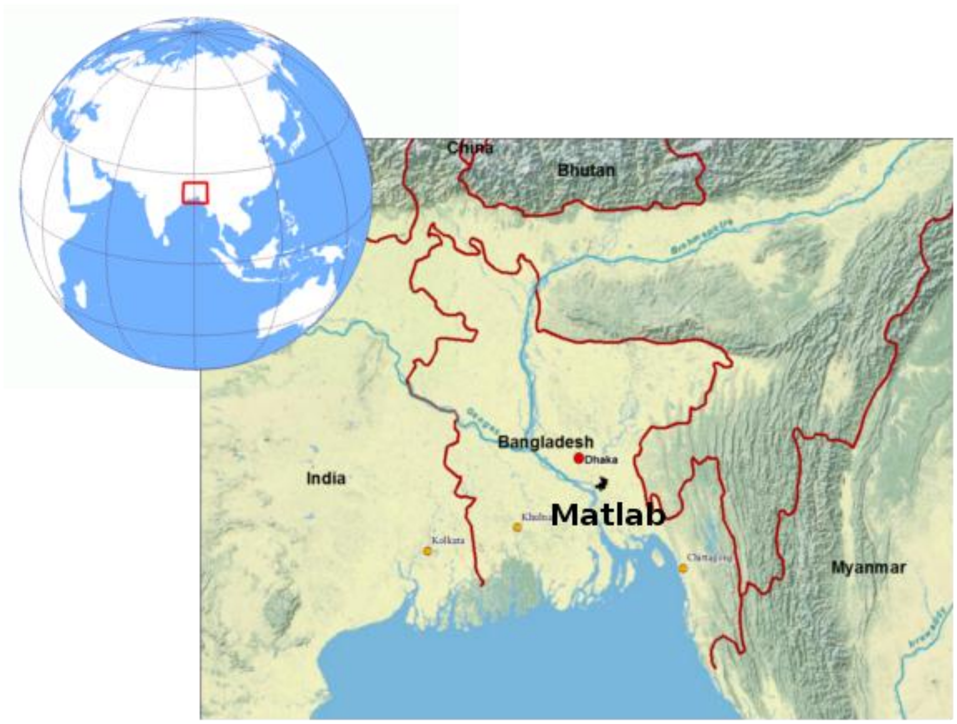

Figure 1: Matlab study area.

Since the 1960s the International Center for Diarrheal Disease Research, Bangladesh (ICDDR,B) has conducted a detailed surveillance of the population, recording all births, deaths, and migrations as well as conducting periodic censuses. These datasets contain health outcomes, socioeconomic status, and socio-demographic information that can be linked for each individual across years. Records can be further linked to the existing Matlab Geographic Information System (MGIS; M. E. Emch 1999) which contains spatial data of the region including point locations of housing (*baris*) and drinking water wells with measured arsenic concentrations.

Beginning in 1977, ICDDR,B implemented a maternal and child health and family planning (MCH/FP) program in half of Matlab, referred to as the treatment area (Figure 2). The remaining villages form two comparison areas, one in the north and the other in the southwest, adjacent to the treatment area. The population in these areas is also recorded in the health and demographic surveillance system, but received the standard government services and is considered the comparison or control areas. Within the treatment area there is the main Matlab hospital as well as three medical subcenters which provide outpatient care, health and family planning information and are staffed by trained paramedics. In later years, the MCH/FP expanded its services beyond fertility and reproductive health services to include other issues such as lower respiratory infections (Stewart, Fauveau, Parker, Chakraborty, & Kham, 1994).

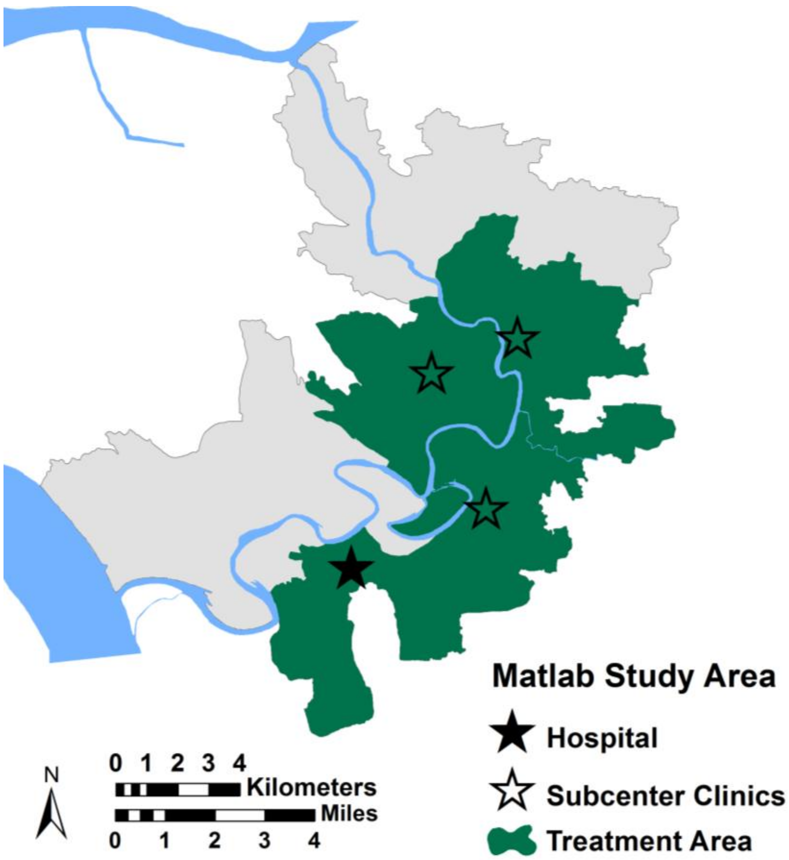

Figure 2: Matlab study area showing locations of medical facilities as well as the treatment area where specialized maternal and child health intervention programs were implemented.

Matlab ALRI Control Program

Beginning in 1989, an ALRI control program was established within the MCH/FP treatment area of Matlab. The program has been described in detail elsewhere (Fauveau, Stewart, Chakraborty, & Khan, 1992; Stewart et al., 1994). The goal was to reduce mortality in children under five through early identification of cases, treatment, and referral for further medical care. As part of the program, community health workers (CHW) were trained to identify ALRI based on a modified WHO case definition using respiratory rates and other visible symptoms such as chest in-drawing (WHO, 1991). ALRI was classified as mild, moderate, or severe. Mild cases were monitored and mothers were given additional health information for supportive care. A moderate case, diagnosed as respiratory rates greater than 50 breathes per minute but without other symptoms, was treated with antibiotics. Between 1989 and 1991, the treatment area was further subdivided into two halves. One treatment sub-area received in-home injections of penicillin by the CHW, while the other half received oral ampicillin syrup at one of the nearby treatment subcenters. No differences in mortality from these two treatment options for moderate pneumonia were detected and all treatment areas later switched to oral co-trimoxazole tablets following the Government of Bangladesh National ARI guidelines (Stewart et al., 1994). Severe pneumonia cases, defined by the presence of chest retractions and other symptoms along with respiratory rates above 50 breathes per minute or any lung infection in children under 1 month old regardless of symptoms, were referred to the ICDDR,B hospital in Matlab where oxygen, intravenous fluids and antibiotics were available.

Previous evaluations of the Matlab ALRI control program have reported it be successful in reducing under 5 mortality by 32 percent in its first two years (Fauveau et al., 1992), and by 54 percent in children less than 2 years (Ali et al., 2001) when compared to the comparison areas. Note however that within the MCH/FP area, families also received other health interventions such as vaccinations, general referral and treatment at subcenter clinics, and family planning services which reduced the number of children and crowding, and promoted breastfeeding. These interventions appear to also improve ALRI mortality as prior to the start of the ALRI-specific program, the treatment area already reported significantly lower mortality rates (Fauveau et al., 1992). After the start of the ALRI control program, the treatment area experienced a further decline in mortality while the comparison areas did not which is evidence of a benefit from the specific intervention program beyond the non-specific benefits of the MCH/FP programs (Fauveau et al., 1992; Stewart et al., 1994).

Acute Lower Respiratory Infections

Acute lower respiratory infections or ALRI are the general name used to identify various bacterial and viral infections in the lower airways (trachea, bronchi, bronchioles, and alveolar sacs). The lungs are particularly vulnerable to infections due to the close contact with the outside environment during respiration and a variety of pathogens can cause infections in the lower respiratory tract (Mizgerd, 2008). Most deaths due to ALRI are attributed to pneumonia or bronchiolitis (Simoes et al., 2006). While the specific pathogen causing a lower respiratory infection is rarely known, common bacterial causes for these diseases are *Streptococcus pneumoniae* ("pneumococcus") or *Haemophilus influenzae* type b (Hib), and viral infections are often caused by respiratory syncytial virus (RSV; Simoes et al. 2006). These microbes can be introduced to the lower lungs in several ways: they may enter during normal breathing; they may be present from a prior upper respiratory infection and migrate down the respiratory tract; or they can be aspirated into the lungs from the esophagus (van der Poll & Opal, 2009). Once in the respiratory tract the innate immune response of the lungs and respiratory system try to eliminate foreign microbes through a combination of defenses starting with mucus production, cilia action, and cough reflexes (Mizgerd, 2006). Failing these first-line, anatomical defenses, the foreign infectious agents are attacked by macrophages in the alveolar sacs. However this action initiates an inflammatory response that draws additional leukocytes (white blood cells), plasma and other fluids to the lungs causing the impaired ventilation and lung function most commonly associated with ALRI diagnoses (Lanata & Black, 2008; Mizgerd, 2008; van der Poll & Opal, 2009). In addition to labored breathing and coughs, ALRI typically present with general symptoms of infection including fever, chills, and malaise (van der Poll & Opal, 2009). In western countries diagnosis of ALRI is usually accomplished by radiography and/or bacterial culture. However, in developing countries and other resource-limited environments, the WHO recommends a case detection strategy based on respiratory rates (WHO, 1991).

Burden, Risk Factors, and Prevention

ALRI cause a greater burden of disease in the world than HIV/AIDS, cancers, diarrheal diseases, or malaria, and despite being a common and persistent threat to health and life around the world, lung infections receive disproportionately less attention and research than other diseases (Mizgerd, 2006). Pneumonia remains the leading cause of death in children under 5 years, and it is estimated to be responsible for 18 percent of all deaths or approximately 1.5 million deaths in 2008 (R. E. Black et al., 2010). Geographically, childhood ALRI deaths are most common in South and Southeast Asia, causing 21 percent of all deaths in children under 5 years (R. E. Black et al., 2010). In Bangladesh there is a high burden of ALRI in rural areas with between 0.23 and 0.47 episodes per child-year (Arifeen et al., 2009; Zaman et al., 1997). ALRI have been found to be the leading cause of child mortality (Spika et al., 1989) and over 25,000 deaths in Bangladesh each year may be due to pneumonia (R. E. Black et al., 2010). Within the study area of Matlab, ALRI are still the first or second leading cause of death from communicable diseases with a mortality rate in 2008 of 7.29 per 1,000 males and 11.39 per 1,000 females within the treatment area (ICDDRB, 2010).

The burden of ALRI falls mostly on the very young and very old and, in general, persons with suppressed or weakened immune systems (Mizgerd, 2006). Other known risk factors for lung infections are related to impaired immune systems and exposure to pathogens or lung irritants. Malnutrition, low birth weight, limited breastfeeding, and lack of child immunizations, all of which limit the development of the immune system, have been associated with increased ALRI risks. Exposure to indoor air pollution from cooking fires or smoking, as well as crowding in indoors spaces also appear to lead to more ALRI because of lung irritation exacerbating symptoms or increased circulation of pathogens (Rudan et al., 2008). Therefore most prevention strategies typically focus on immunization and improved nutrition in children in order to improve immune status (Sazawal & Black, 2003). Vaccines, particularly for measles and Hib, have been found to be successful in preventing forms of ALRI. But these vaccines must be given in very early infancy because young children are at such high risk of infection (Arifeen et al., 2009). The wide variety of current and emerging pathogens with the potential to cause ALRI, is an additional hindrance to any vaccination campaign (Mizgerd, 2006) so that even with newer vaccines and adequate coverage a significant number of new cases are likely to develop (Arifeen et al., 2009). Given these challenges, a case management strategy reducing the severity of disease and preventing mortality is a more feasible option in many areas of the world (Mizgerd, 2006; Sazawal & Black, 2003). Community-based programs designed to provide early case detection and treatment with antibiotics, such as the one established in Matlab, have proven successful in reducing child mortality from ALRI, particularly in less-developed areas. In their review and meta-analysis of nine studies of pneumonia case management programs, Sazawal and Black (2003) report an average 36 percent reduction in mortality for children under 5 years old.

Arsenic

Inorganic arsenic is a significant environmental factor affecting human health, and these problems are particularly acute in Bangladesh where between 35 and 77 million people have been exposed to high concentrations of arsenic in their drinking water over the years (Smith et al., 2000). Although a naturally occurring, semi-metallic element in the alluvial sediments of the Bengal Basin region, arsenic is a danger to human health when large quantities are ingested through contaminated food and drinking water (Nickson, Mcarthur, Ravenscroft, Burgess, & Ahmed, 2000). Arsenic is a potent toxin and carcinogen as long-term exposure from contaminated drinking water often leads to bladder, kidney, lung, and skin cancers (National Research Council, 2001; Smith, Goycolea, Haque, & Biggs, 1998). Other health complications include lesions and discoloration of the skin and higher risk of cardiovascular and infectious diseases (Argos et al., 2010; Smith et al., 2000; Sohel et al., 2009).

Biological Effects of Arsenic Exposure

Arsenic can occur in different chemical states when dissolved in water – arsenite (trivalent arsenic or AsIII) and arsenate (pentavalent arsenice or AsV; Rossman 2007). These different oxidation states affect the general level and mechanism of arsenic's toxicity (Oremland & Stolz, 2003). Arsenate typically affects mitochondrial metabolism and ATP molecules which are responsible for cellular metabolism and energy, while arsenite has a more complex process affecting cellular functions such as signaling proteins and even damaging DNA (Rossman, 2007; Vahter & Concha, 2001). Regardless of valent state, after arsenic is ingested, it is absorbed in the gastrointestinal tract and carried primarily to the liver and kidneys where, as well as by cells in the bloodstream, the first steps of metabolism occur (Rossman, 2007). Multiple stages of metabolism by methylation (attaching a methyl group, CH_3) and redox reactions are necessary to transform arsenic into the more readily excreted compounds monomethyarsonic acid (MMA) and dimethylarsinic acid (DMA; Vahter and Concha 2001; Vahter 2002). Arsenic is then excreted from the body primarily in urine but also in hair and nail tissues (Rossman, 2007). The end result of the methylation process transforming arsenic is often considered a detoxifying process because MMA and DMA are less reactive and more readily excreted substances; however, the intermediate metabolites that are created in the process (specifically MMA III and DMA III) are, in fact, more potent toxins and more reactive to the human body than other forms of arsenic (Vahter, 2002). The toxic effects of chronic arsenic exposures appear to derive from repeated interactions of these intermediate arsenic forms with tissues while circulating in the body during metabolism. The process of methylation may not always reach completion and therefore some of the ingested inorganic arsenic remains as the significantly more toxic MMA III. In reviewing multiple epidemiologic studies of urinary arsenic concentration, Smith and Steinmaus (2010) find that disease risk increases with higher proportions of MMA measured in urine. Therefore, individual differences in arsenic metabolism rates may affect disease outcomes (Vahter & Concha, 2001).

The specific toxic effects of arsenic operate on the cellular and molecular level and growing evidence from *in vitro* tests, animal models, and epidemiologic studies suggests that the immune system is a primary target (Kozul, Hampton, et al., 2009; Lemarie, Morzadec, Bourdonnay, Fardel, & Vernhet, 2006). Arsenic disrupts both the innate and acquired immune responses by limiting functions of leukocytes (white blood cells) as well as the proteins used in contact-dependent cellular communication. Specifically, arsenic has been found to affect lymphocytes, particularly helper T cells (T_h) and Natural Killer (NK) cells, and their secretion of a protein needed to stimulate further immune response (Andrew et al., 2007, 2008; Soto-Peña et al., 2006). Fewer macrophages are available in arsenic-exposed individuals and these cells, which are responsible for destroying pathogens, showed decreased adhesion and ingestion of foreign material (phagocytosis) as well as morphological changes in their structure (Banerjee et al., 2009; Lemarie et al., 2006). Consequently, cellular damage becomes more difficult for the body to repair following prolonged arsenic exposure (C. E. Olsen et al., 2008), and these changes in the immune system and cellular functions directly limit the body's ability to provide a sufficient and coordinated response to inflammation and infection. Evidence of negative impacts on the immune system classifies arsenic as an immunotoxicant (Selgrade, 2007). Environmental exposure to this class of toxins, which includes chemicals and metals, may partly explain burdens of infectious diseases (Selgrade, 2007; Winans, Humble, & Lawrence, 2011). Moreover, cells within the lungs seem to be particularly susceptible to the negative effects of arsenic. Sherwood et al. (2011) report that arsenic prevents epithelial cells within the lung from sending and receiving signals necessary to coordinate wound closure, cililal beat, and mucus secretion as part of the innate immune response. A study of mice exposed to arsenic-contaminated drinking water found that the genes needed for innate immune response were significantly altered in lung tissue (Kozul, Hampton, et al., 2009). Such damage to lung tissue, coupled with a disruption in the inflammatory and immune responses could contribute to increased risks of respiratory infections (Kozul, Hampton, et al., 2009; Kozul, Ely, Enelow, & Hamilton, 2009).

Geological Basis for Groundwater Arsenic

While once thought to be anthropogenic pollution, the source of arsenic in the Bengal Basin Region is now considered to be the result of natural geologic conditions (Smedley & Kinniburgh, 2002). Arsenic typically makes up about 1.8 ppm of the earth's crust, and while slightly higher proportions have been reported in the sediments around Bangladesh, the area is not unique (Ravenscroft, Burgess, Ahmed, Burren, & Perrin, 2005). The unusual magnitude of the arsenic problem in this region comes from the unique conditions in the aquifers allowing for widespread release of arsenic into the groundwater.

Arsenic in the sediments and groundwater around the Ganges-Brahmaputra-Megna River systems originated in the Himalaya of India and Nepal and through a process of weathering and transport arrived in the lower floodplains starting during the Quaternary and earlier periods more than 20,000 years BP (Ravenscroft et al., 2005; Smedley & Kinniburgh, 2002). These deposition processes occurred in multiple steps over thousands of years, as described by Ravenscroft et al. (2005) and Nickson et al. (2000). Weathering of upland area rocks provided the first mobilization of arsenic which was then adsorbed primarily to iron oxyhydroxides (FeOOH). These oxyhydroxides coated fine-grained sediments and were transported via rivers and streams and collected on the delta and floodplains of the Bengal Basin.

The fate of these sediments and the arsenic-containing oxyhydroxide deposits appears to be strongly linked to the paleo-climate and hydrology of the period when they were deposited. During the earlier Plio-Pleistocene, the cooler climate and substantially lower sea level allowed for oxidation of the iron in the floodplain sediments preventing the release of arsenic while rivers and rainfall were also able to carry away the arsenic-carrying sediments (Nickson et al., 2000; Ravenscroft et al., 2005). The oxidized sediments from this period appear yellow or brownish-red in color and aquifers found in these older layers are generally free of arsenic (Hoque, Burgess, Shamsudduha, & Ahmed, 2011). Rising sea levels following the last glacial maximum, approximately 11,000 years BP, redeposited more arsenic-containing sediments along the floodplains and former valleys while the rising water table and low hydraulic gradient limited subsurface flow and removal of arsenic. The warmer climate of the Holocene also enabled dense swamps to grow in the lowland areas of the Bengal Basin Region providing a source of rich organic matter (Ravenscroft et al., 2005). The presence of organic matter (peat) buried with the layers of arsenic-containing sediments during this time period is a key element in the process releasing arsenic into the groundwater. Over thousands of years, microbial activity consumed the buried organic material creating an oxygen- and nitrate-limited environment (Nickson et al., 2000). These decomposition actions created strongly reducing conditions in the aquifers of this period, thus causing iron oxyhydroxides in the sediments to dissolve, desorbing arsenic into the groundwater (Nickson et al., 2000; Ravenscroft et al., 2005). Microbes may also play additional roles in releasing and cycling arsenic into the groundwater by directly breaking down the sediments containing arsenic (Oremland & Stolz, 2003). The dissolved arsenic is then free to be brought to the surface through groundwater removal while continued low hydraulic gradients and long residence times in the Bengal Basin Region prevent arsenic from being flushed from these aquifers.

Tubewell Use and Arsenic Contamination

In Bangladesh, exposure to groundwater contaminated with high concentrations of arsenic is the unintended consequence of attempts to alleviate morbidity and mortality from diarrheal and other water-borne diseases. Beginning in the early 1970s UNICEF supported the Bangladeshi government in installing about 1 million tubewells for drinking water. The program expanded further in the 1980s with approximately 3 million more wells by providing affordable loans to individuals to privately install their own wells (Bearak, 1998). The wells that were installed, also called tubewells, consist of a narrow (5-10 cm) metal or PVC plastic pipe. The soft soil allowed many wells to be installed manually and most were initially sunk into the earth only 10 to 40 meters deep, reaching the highest water table (M. Black, 1990). The tubes are then attached to hand-operated pumps to draw water to the surface. Some deeper wells may use a small engine to pump water, though these wells are primarily used for irrigation rather than drinking water (Paul & De, 2000). Many of the aquifers accessed by these new wells, however, contained dissolved arsenic. At the time these wells were being installed, arsenic was not known to contaminate groundwater in the Bengal Basin region or in other similar geologic formations so testing for it was uncommon. Arsenic-contaminated water has no distinguishing color, smell, or taste, so it was not until the 1980s and 1990s that health problems in the region, beginning with skin lesions, were linked to arsenic exposure, that wells were tested, and the extent of the problem became known (Smith et al., 2000).

Variability in Arsenic Levels

In the study area of Matlab, 62 percent of the over 13,000 functioning tubewells tested in 2002-2003 had arsenic levels above the Bangladeshi-government-recommended level of 50 µg/L (approximately 50 parts per billion) and 9% have concentrations above 500 µg/L (ICDDRB, 2004). The U.S. EPA and the World Health Organization recommend a limit of only 10 µg/L. Owing to the complex subsurface hydrogeological and chemical conditions necessary for arsenic to become mobilized into the groundwater there are substantial variations in the concentrations measured. A national sample of tubewells surveyed by the British Geologic Survey (BGS) and the Bangladesh Department of Public Health and Engineering (DPHE) found significant variation in arsenic levels; they report a range of arsenic concentrations from 0.25 µg/L to over 2,000 µg/L with the highest values found in the south and southeast regions of the country near the Chandpur District, the area containing Matlab (BGS & DPHE, 2001). Within the larger patterns there is tremendous local-scale variation. The BGS-DPHE survey used a variogram analysis to estimate that 40% of the spatial variation in arsenic levels occurs within 2 km, indicating rapid changes in arsenic levels between neighboring tubewells.

These local variations are partially explained by differences in tubewell and aquifer depths. As discussed above, arsenic in groundwater is primarily found in relatively young (Holocene) sedimentary layers which contain the necessary combination of arsenic-bearing sediments, dissolved organic matter, and bacteria. The location and depth of these sediments and aquifers are not uniform across the country but are typically 20 to 50 meters below ground. Productive aquifers are also found in the deeper Plio-Pleistocene layers, and these aquifers appear to be hydraulically separated from one another by layers of clay (Hoque et al., 2011; Smedley & Kinniburgh, 2002). Generally, aquifers found deeper than 130m are older and are reported to have low arsenic concentrations because of their different geologic conditions (BGS & DPHE, 2001; Dhar et al., 2008) while shallow aquifers tend to have higher arsenic levels but also higher variability in arsenic concentration (van Geen et al., 2003; Ravenscroft et al., 2005). Additionally, hand-dug wells shallower than 5 m are also frequently safe from high arsenic levels due to frequent flushing (Ravenscroft et al., 2005). In the past, when arsenic contamination was not known to be a problem, most tubewells were only installed approximately 50 meters deep to reach the first viable aquifers (M. Black, 1990). While this decision reduced drilling and installation costs, the shallow wells tapped relatively young aquifers with some of the highest levels of arsenic due to this relationship between aquifer depth, age, and hydrogeologic conditions. It has been suggested, though, that the high spatial variability in arsenic concentrations can be exploited to partially mitigate exposure by encouraging villages or households to switch to nearby arsenic-free wells (van Geen et al., 2002).

There is great concern that there may also be temporal trends in arsenic levels due to changes in sediments, release/uptake of arsenic, and flushing of aquifers; however, the current evidence generally supports the conclusion that arsenic levels are stable. The BGS-DPHE study found very limited evidence for temporal trends in arsenic concentrations over 12 months (2001). Longer time series data monitoring arsenic levels are very rare in the Bengal Basin. From the few published studies available (none covering a period longer than 3 years) there are conflicting reports about a long-term increasing trend (see table 1 in Cheng et al. 2006; Peter Ravenscroft, Howarth, and John M Mcarthur 2006). Cheng et al. (2005) found little change in arsenic concentration from baseline measurements over 3 years of biweekly monitoring. But in other areas arsenic concentrations have been found to be marginally higher in older wells, leading some to suggest a possible increase in arsenic levels over the long-term (van Geen et al., 2003; Ravenscroft et al., 2006). Well failure that begins drawing water from shallower aquifers and pumping-induced movement of high-arsenic water that contaminants previously safe wells have been offered as possible explanations for the association between well age and arsenic level (Cheng et al., 2006; Ravenscroft et al., 2005). However, one pattern that does emerge from some temporal studies of arsenic levels is the existence of a seasonal cycle that varies arsenic concentrations within a well over the year as groundwater levels fluctuate and aquifers are filled by heavy monsoon rains. One report from Bangladesh found minimum concentrations of arsenic occur primarily during the summer months while peaks occur during the monsoon (Savarimuthu et al., 2006). A similar pattern of wet season peaks was found in Vietnam (Berg et al., 2001) though others have not found any significant seasonal pattern (Dhar et al., 2008).

There has been relatively little monitoring of arsenic concentrations, though temporal trends are particularly important for epidemiological studies estimating past exposures and could cause over- or under-estimates of cumulative doses of arsenic. The evidence for long-term trends (either increase or decreasing) in arsenic remains in favor of stable levels, though more work is needed. The small spatial areas covered in most monitoring studies also limits their generalizability to other areas, particularly given the evidence already showing large spatial variations in arsenic levels (Sengupta et al., 2006). In the present study, any seasonal cycle likely has little influence, and, in the short and medium term, measured arsenic levels are assumed to be generally stable.

Social Patterns of Arsenic Exposure

In addition to the hydrogeological factors that influence arsenic levels across the region, exposure to that arsenic varies with the population distribution as well as socioeconomic characteristics. The program encouraging people in Bangladesh to switch to tubewells was widely successful with at least 95 percent of all rural resident adopting tubewells as their primary source of drinking water by the early 1990s (Caldwell et al., 2003). However tubewells were not installed or adopted uniformly across the area. Higher status households were early adopters of the new water source and tubewell ownership was considered a status symbol (M. Black, 1990; M. Rahman et al., 2006). The program to install tubewells was deliberately designed to avoid bureaucracy by being decentralized as local villages and individuals paid for the installation and maintenance of the wells (M. Black, 1990). Guidelines for the selection of new wells sites were established to try to ensure an equitable distribution and access to tubewells. However, in many instances, the wealthiest household in an area was assigned ‘caretaker’ status of a new well and, thus, that family had more influence on the site selection and access decisions (M. Black, 1990). Knowledge of arsenic contamination and its health risks were slow to spread in rural areas and to lower status individuals (Paul, 2004). More recent surveys of arsenic exposure have found evidence that lower socioeconomic status households are more likely to be exposed to high levels of arsenic (Khan, Aklimunnessa, Kabir, & Mori, 2007). Higher status families are likely able to leverage their resources or those of their social networks to install the more costly, but safer, deep tubewells or to gain access to an alternative, safer well.

Diseases Associated with Arsenic Exposure

Given the damaging effects of arsenic to the immune system, as well as to cells and DNA, it is not surprising that wide-ranging health problems have been associated with prolonged human exposures to contaminated drinking water. Symptoms of chronic arsenic poisoning (arsenicosis) typically appear after 5 to 10 years or more of exposure and are classically presented in skin discoloration (hyper- or hypo-pigmentation) on the arms and chest, and nodules on the hands and feet known as keratosis (Guha Mazumder, 2008; Smith et al., 2000). These skin conditions can lead to further infection and gangrene. Other effects of arsenic are less visible, but more deadly. Higher cumulative doses of arsenic increases all-cause mortality (Sohel et al., 2009), and chronic exposure to elevated levels is also associated with increased risks of specific conditions including cancers to the skin, lungs, and bladder, as well as increased risk of other diseases such as cardiovascular disease and diabetes (Smith et al., 2000). Importantly, arsenic has been found to be transmitted across the placenta (Concha, Vogler, Lezcano, Nermell, & Vahter, 1998) where it disrupts the immune system development of the fetus (Ahmed et al., 2011). Because of this transmission, arsenic also affects birth outcomes and the early life health and survival of infants. In utero exposures have been found to increase risks of low birthweight, spontaneous abortions, and later cognitive development (Bloom, Fitzgerald, Kim, Neamtiu, & Gurzau, 2010; Vahter, 2009).

Of particular interest to the present study is the growing epidemiologic evidence that ingested arsenic is also associated with higher risks of non-malignant lung diseases and respiratory symptoms (Guha Mazumder et al., 2000; Guha Mazumder, 2007; Milton, Asan, Rahman, & Rahman, 2001). Lung disease is typically associated with inhalation exposures (eg. cigarettes or asbestos); however, Smith and colleagues (2009) report that the effects of arsenic on the lungs appear to be the same whether it is inhaled or ingested through contaminated water which they claim makes it a uniquely toxic substance to the lungs. Arsenic exposure appears to increase symptoms such as coughs and decrease volume and flow of air that can be moved in the lungs (Smith & Steinmaus, 2009). Infants exposed in utero to arsenic have higher rates of acute respiratory infections (Raqib et al., 2009). In a study based in India, von Ehrenstein et al. (2005) found a significant decrease in such lung function in men who had been exposed to arsenic contaminated drinking water, and, similarly, Milton et al. (2001) also found higher prevalence rates of chronic bronchitis and coughs among Bangladeshis who were exposed to arsenic through their drinking water. Mortality from lung diseases also seems to increase. In the years following a spike in arsenic in water supplies in northern Chile, adult mortality from pulmonary tuberculosis increased (Smith et al., 2011) and young adults who were exposed in childhood experienced higher mortality rates from bronchiectasis (Smith et al., 2006).

Matlab has been the site of a significant amount of the research on the health effects of arsenic. The drinking water contamination there is causing excess adult and infant mortality and producing a variety of chronic ailments in the population (Yunus, Sohel, Hore, & Rahman, 2011).

Limitations of Previous Research

Past studies of both arsenic poisoning and ALRI have typically focused on individual-level outcomes. These studies have emphasized the biological consequences of arsenic exposure or the specific symptoms and disease agents of ALRI. While this largely biomedical approach can be useful for clinical or curative studies, there is a need to consider the broader determinants of health (Loeb, 2003). There is evidence that community-level socioeconomic status can also play a role in lung infection risk (Cohen, 1999), and ecological analyses can help understand what factors beyond individual-level risk behaviors can be driving spatial patterns in ALRI (Crighton, Elliott, Moineddin, Kanaroglou, & Upshur, 2007). A similar pattern of individual-level focus can be seen in the arsenicosis literature. While some studies have examined socioeconomic (Mosler, Blöchliger, & Inauen, 2010) and legal (Atkins, Hassan, & Dunn, 2006) dimensions of arsenic exposure, there has been relatively less work (but for an exception see Sohel et al., 2010) on the spatial variation in arsenic exposure beyond regional patterns (Paul & De, 2000), geostatistical interpolation of potential exposure (Hassan, Atkins, & Dunn, 2003), or geological variation in arsenic concentrations (van Geen et al., 2003). A broader and more holistic view of the determinants of health, and one that incorporates spatial dimensions is needed for a more complete understanding of both ALRI and arsenic and for designing and evaluating place and context-specific intervention strategies.

Previous studies have also been limited methodologically and have not made full use of spatial statistical methods developed in geography. Prior studies of spatial patterns of disease in Matlab have been limited to exploratory visualizations of patterns. These studies have primarily used a spatial filtering technique that presents a smoothed and continuous surface of risk factors or disease rates (Ali, Emch, Donnay, Yunus, & Sack, 2002). In the study motivating much of this work, Ali et al. (2001) employ such a spatial smoothing technique to illustrate areas of higher ALRI mortality rates in Matlab. However they define high risk areas at a natural break in the distribution and they cannot test whether the local areas of elevated mortality have statistical significance. Some studies have also attempted to employ local clustering tests. For example, Sohel et al. (2010) used a spatial scan statistic to identify local clusters of fetal loss and infant death. They also used a local Moran's I statistic to find local areas of elevated arsenic levels. However, they stopped short of combining their mortality and arsenic analyses. They did not adjust the scan statistic to test whether the clusters were due to the presence of elevated arsenic levels and, thus, their results were limited to a qualitative, visual assessment that a cluster of higher rates occurred in areas that have higher levels of arsenic. By controlling for risk factors and comparing outcomes this study could have tested whether arsenic was in fact responsible for the cluster. Methods exist to allow for cluster detection tests to be adjusted for a continuous measure by means of a regression equation to incorporate environmental risk factors such as arsenic; however, this technique is underutilized throughout the literature, with only two examples to date (Huang, Kulldorff, & Gregorio, 2007; Klassen, Kulldorff, & Curriero, 2005).

Research Approach

This research is guided by the theory of disease ecology, one major theoretical framework within the sub-discipline of health and medical geography. As part of the tradition of human-environment interaction studies in geography, disease ecology views health as the result of interactions between environmental, social/behavioral, and biological systems in space and time (Meade and Earickson, 2000). In this research I consider ALRI to be the result of these three broad, interacting factors which will also incorporate the effects of the ALRI intervention program and arsenic exposures. The analytical approach applied uses a combination of spatial and non-spatial methods in order to understand what factors are influencing the spatial patterns of this disease in Matlab. Together, this theoretical framework and methodological approach seek to address the limitations of previous studies.

Theoretical Framework

Disease ecology focuses on the ways in which human behavior interacts with the total environment in specific contexts to produce health and disease in a population (Meade and Earickson, 2000). The core concepts of disease ecology can be traced to Jacques May (1958) and Rene Dubos (1965) who each articulated similar ideas that health is a result of disease agents interacting with a combination of multiple cultural and environmental factors. From its beginning, disease ecology and medical geography have emphasized human-environment interactions (Hunter, 1974). Meade (1977) felt that May's early work was largely static and descriptive. She further developed disease ecology and reorganized the fundamental concepts using a triad of broad categories from human ecology: population, habitat, and environment. Population factors include the biological and demographic aspects of a population (age, gender, genetics, etc.). Habitat is broadly understood as the environment, both built and natural. Behavior incorporates the social and cultural factors and larger political systems which influence society. These factors naturally interact with each other and over time attempt to maintain a state of dynamic equilibrium (Mayer, 2000). When this balance is disturbed out of equilibrium then changes in health are the consequence. Each of the vertices of the triangle of human ecology is expressed differently over the landscape and they combine to generate spatial variations in health and disease. In this way, specific contexts and place are central ideas to medical geography.

Note that while Meade's triangle of disease ecology at first appears similar the classic epidemiological triangle of agent, host, environment, it differs in important ways. In disease ecology there is less emphasis on individuals as biological entities. This approach does not deny the importance of biological disease agents, but disease ecology sees them, along with psycho-social, political factors, and environmental characteristics, among others, as outside stimuli or insults that are produced by and interact with the state of health formed by the population, habitat and behavior (Mayer & Meade, 1994). This theoretical approach requires a broader and more dynamic view of health, moving toward one in which health is seen as a varying property not simply the absence of disease or specific pathogen (Meade & Earickson, 2000).

Disease Ecology of Acute Lower Respiratory Infections

Figure 3 shows graphically how the disease ecology approach can be applied to this research on ALRI. The known risk factors for respiratory infections, described above, are fit into the population, habitat, and behavior categories. Characteristics of the population include the presence of young children who are at risk as well as factors affecting immune status such as nutrition. Habitat factors include crowding which can increase exposure and transmission of infectious diseases as well as the healthcare “environment” which encompasses the presence and access to treatment centers. Attitudes, beliefs, and social norms surrounding care seeking and knowledge of the symptoms and supportive care needed for respiratory infections are included in the behavioral category.

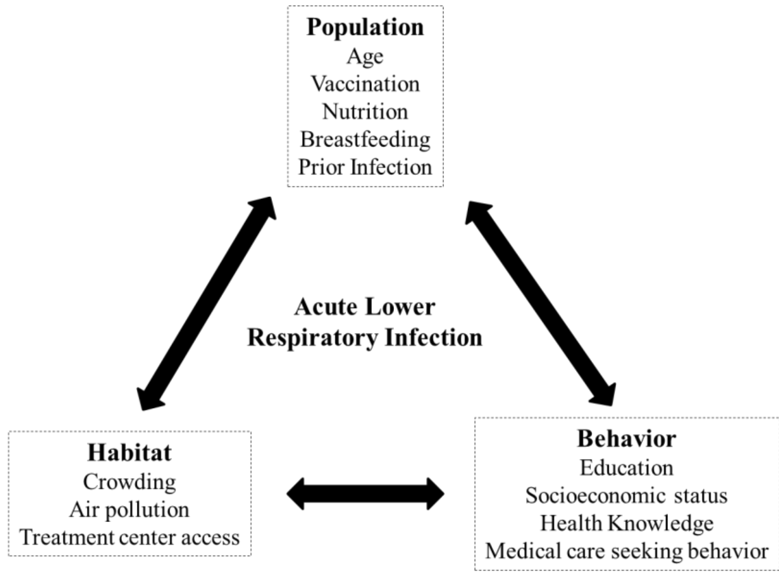

Figure 3: Triangle of human ecology applied to Acute Lower Respiratory Infections.

In the context of Matlab, Bangladesh there are two added stimuli affecting this ecology (Figure 4). First is the ALRI control program which existed in a geographically limited area of Matlab. As part of the MCH/FP, this human response to the high rates of morbidity and mortality attempted to alleviate these infections by influencing all three vertices. The program targeted behaviors by increasing caregivers' knowledge to identify ALRI early and provide appropriate supportive care. The habitat was altered by increasing access to care through visits by community health workers and referrals to specialized treatment centers. The modifications to behaviors and habitat affect the population by reducing prior infections, but emphasizing vaccines as part of the MCH/FP and making antibiotic treatment of respiratory infections more available further improved population health. The second insult of interest is arsenic-contaminated drinking water (Figure 5). This stimulus again began with an intentional reaction to the problem of endemic diarrheal disease. The tubewell program attempted to modify the region's disease ecology by introducing new technology and changing behaviors about the choice of water source. Arsenic in the natural environment is a feature of the habitat, and the changes introduced by the tubewell produced the opportunity for the population to be exposed. The end result is that arsenic exposure produces biological impacts in the population with damage to cells and the immune system. As noted earlier, policies and socioeconomic status varied who and which areas gained early access to these wells. There is also a clear spatial component to this exposure as arsenic varies across the area. When applying a disease ecology framework to Matlab, human-environment interactions and a spatial approach are necessary to understand the multiple influences on ALRI.

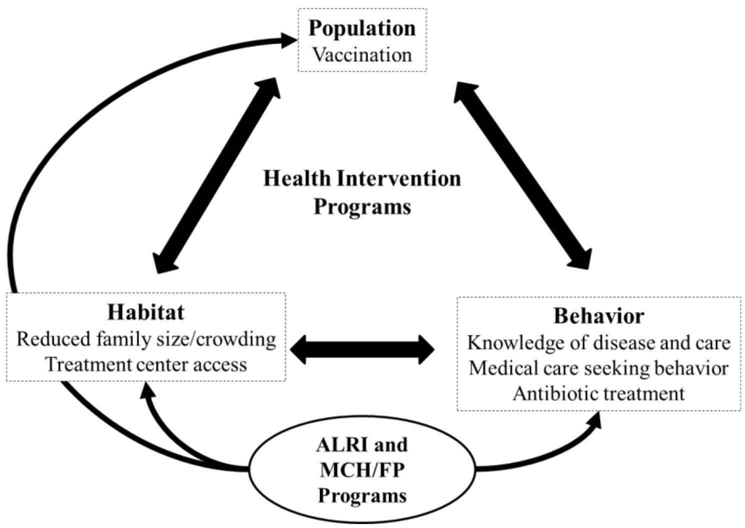

Figure 4: Theorized effects of the Maternal and child health and family planning (MCH/FP) and acute lower respiratory infection (ALRI) intervention programs on the disease ecology of Matlab, Bangladesh.

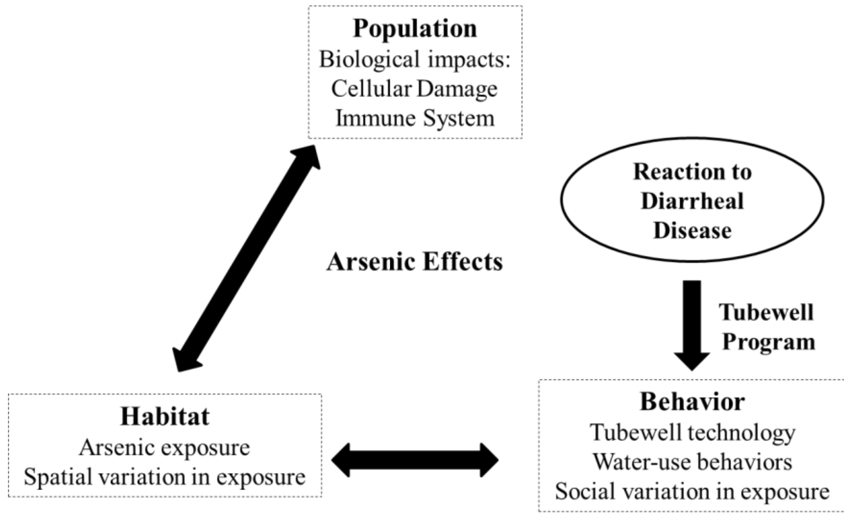

Figure 5: Additional stimuli of arsenic exposure on the disease ecology in Matlab, Bangladesh.

Methodological Approach

In order to understand the spatial patterns of health, disease ecology incorporates methods both from within and outside the discipline of geography. Methods developed in medical geography are gaining wider use in in epidemiology, and spatial analysis of health outcomes is a highly active area of research (Glass, 2000). Exploratory disease mapping and spatial analytical methods have long been a mainstay of medical geographers, but newer spatial methodologies allow us to move beyond simply mapping rates or observations of diseases. The methods used in this research fall clearly into the category of spatial analysis and quantitative geography and include regression analyses of bari-level mortality rates which are then used to explore the spatial patterns and clusters of ALRI in a spatial scan statistic.

The regression techniques used in this research are zero-augmented models. These models are designed for counts of events and are appropriate for rare outcomes which produces excess zeroes (Cameron & Trivedi, 1998; Lambert, 1992; Loeys, Moerkerke, De Smet, & Buysse, 2012). These generalized linear models are two-part models made up of a binomial model predicting the probability of zero events beyond what would be expected by a Poisson or negative binomial model. The second part of the model then uses a Poisson or other similar probability distribution to model the observed counts. Together these two parts describe the overall distribution of events while reflecting the inflated probability of not reporting any events. By including an offset population, rates accounting for differences in the population at risk can also be calculated from these models. While these models have been previously used to model health outcomes (Carrel et al., 2011; Carrel, Voss, Streatfield, & Yunus, 2010), they have never been combined with cluster detection techniques.

Cluster detection includes a variety of spatial statistical methods designed to detect the presence of unusual spatial and/or temporal groupings of events. There remain significant concerns about tests of disease clusters, beginning with how one defines both the study area and what constitutes an unexpected cluster to ensure an unbiased estimate (Elliott & Wakefield, 2001). Researchers are cautioned that clustering tests also can be difficult to interpret and link to a single social or biological pathway that can be then linked back to the disease (Elliott & Wakefield, 2001). Moreover, the statistical power of the tests can be weakened when the tests are applied to a pre-supposed cluster or without first constructing a hypothesis for possible etiology, *a priori* (Elliott & Wakefield, 2001; S. F. Olsen, Martuzzi, & Elliott, 1996). While these concerns are valid, cluster analysis is frequently applied in earlier stages of an investigation. If clusters are detected then the next phase should include a detailed, local study to test possible mechanisms and hypothesized risk factors that could cause the clustering (Wakefield, Kelsall, & Morris, 2001). In this study, I employ a local cluster analysis using the spatial scan statistic. The spatial scan statistics developed by Kulldorff (1997) addresses several of these concerns: it tests all areas equally and does not presuppose the existence of a cluster;

statistical significance is adjusted for multiple testing; and the test can be adjusted for known risk factors (Kulldorff, Feuer, Miller, & Freedman, 1997). The cluster analysis uses a zero-inflated regression to adjust for possible risk factors and the resulting cluster analysis is interpreted, as Wakefield and colleagues, as suggestive for future studies.

Structure of the Report

Following this background and introductory chapter, this thesis continues with two, self-contained research papers designed to answer my research questions examining the spatial patterns of child mortality from respiratory infections and the possible link to arsenic exposure. The first paper focused on the developing the zero-inflated regression model of mortality rates and is structured for a public health journal such as the *American Journal of Epidemiology*. The second paper deals with the cluster analysis and is aimed for a geography journal such as *Health & Place*. As the results of the first paper inform the second, the two chapters are joined by a bridge chapter. Finally, this report concludes with a discussion of the research conducted, the broader impacts of the work, and possible future avenues of research.

Paper 1: Effects of Health Intervention Programs and Arsenic Exposure on Child Mortality from Acute Lower Respiratory Infections in Rural Bangladesh

For submission to the *American Journal of Epidemiology*

Abstract

This paper examines the effectiveness of a community-based treatment program on acute lower respiratory infection (ALRI) mortality in rural Bangladesh. Exposure to inorganic arsenic in contaminated drinking water is a widespread health threat in Bangladesh that may increase mortality and negatively affect the lungs. ALRI mortality data were obtained for children under 2 years old from 1989 to 1996 in the Matlab demographic surveillance system. This study period represents the years immediately following the implementation of the ALRI control program yet before contamination of water sources was known. A zero-inflated negative binomial regression model was used to simultaneously estimate mortality rates and the likelihood of no deaths in groups of related households while controlling for social and environmental characteristics, and access to care. The results suggest that the ALRI control program was successful in reducing child mortality while arsenic exposure was only marginally associated with increased mortality. Higher socioeconomic status also significantly reduced mortality rates.

Introduction

Acute lower respiratory infections (ALRI), which include pneumonia as well as bronchiolitis, tracheobronchitis and croup, are the leading cause of childhood morbidity and mortality globally, responsible for 18 percent of all deaths in children under 5 years (R. E. Black et al., 2010; Lanata & Black, 2008). The greatest burden of these diseases is suffered by children in less-developed countries which experience 97 percent of the 156 million new cases of pneumonia each year (Arifeen et al., 2009; Rudan et al., 2008; Spika et al., 1989; Zaman et al., 1997). In Matlab, Bangladesh ALRI remain among the leading causes of death from infectious diseases (ICDDRB, 2010).

Vaccines designed for the two main causes of ALRI, pneumococcus (*Streptococcus pneumonia*) and Hib (*Haemophilus influenzae* type b), are available and they have been found to prevent lung diseases; however, there are significant political, economic, and logistical challenges for distributing vaccines, particularly in developing countries (Girard, Cherian, Pervikov, & Kieny, 2005). Additionally, the wide variety of pathogens capable of causing ALRI hinders vaccine strategies and ensures that a significant number of new cases are likely to develop (Arifeen et al., 2009; Mizgerd, 2006). In the absence of these prevention strategies, community-based intervention programs designed for early ALRI case detection and treatment with antibiotics have been found to be successful in reducing child mortality from ALRI (Sazawal & Black, 2003). In Matlab, Bangladesh, an ALRI intervention program implemented in 1989 and was successful in quickly reducing child morality by over 50 percent (Ali et al., 2001; Fauveau et al., 1992). However, ALRI remains a persistent cause of morbidity and mortality in Matlab and there are local areas of elevated ALRI mortality in children that remain unexplained.

The widespread contamination of drinking water by arsenic in Bangladesh is an additional serious health threat which must also be considered when examining any of the outcomes of the intervention program in Matlab. Inorganic arsenic is a potent toxin that causes wide-ranging health problems as a result of its damaging effects on the immune system (Selgrade, 2007; Soto-Peña et al., 2006). Cells of the lung seem particularly sensitive to arsenic which disrupts the inflammatory response and innate immune system signaling (Kozul, Hampton, et al., 2009; Kozul, Ely, et al., 2009). As the lungs are in close contact with the outside environment and pathogens, these immune response changes could increase the risk of a lower respiratory infection (Kozul, Hampton, et al., 2009). Previous studies have found that exposure to inorganic arsenic in drinking water negatively affects the lungs, increasing mortality from lung cancer (Smith et al., 1998), bronchiectasis (Smith et al., 2006), and tuberculosis (Smith et al., 2011) as well as decreasing lung function (Guha Mazumder, 2007) and increasing susceptibility to lower respiratory infections (Raqib et al., 2009).

The aim of the present study is to evaluate the population-level relationship between arsenic in drinking water and child mortality from ALRI in the context of a health intervention program which specifically targeted this health outcome. Arsenic exposure was not previously considered when evaluating the impact of the ALRI program in Matlab and the possible link between child mortality from ALRI and arsenic exposure has not been well studied.

Materials and Methods

Study Area

This study focuses on the region of Matlab, Bangladesh, a rural region of over 200,000 people located approximately 50 km southeast of the capital city Dhaka in central Bangladesh (Figure 6). Since the 1960s, Matlab has been the site of a comprehensive health and demographic surveillance system organized by the International Center for Diarrheal Disease Research, Bangladesh (ICDDR,B) which has recoded all births, deaths, and migrations as well as conducted periodic censuses in the region. Matlab is divided into two geographic areas for program administration. Approximately half of the population lives in the treatment area and receives specialized maternal and child health and family planning programs, while the other half, which is also recorded in the demographic surveillance system and census, receives only the standard government services and is considered the comparison area. The population of Matlab lives in patrilineally-related groups of housing units known as *baris*, which are the unit of analysis for this study.

ALRI Control Program

Beginning in 1989, ICDDR,B initiated a community-based ALRI control program within the treatment area of Matlab (Fauveau et al., 1992; Stewart et al., 1994). The goal of the ALRI program was to reduce mortality in children under 5 years through early identification of respiratory infection cases. Treatment options, in order of case severity, included: supportive care and monitoring, outpatient treatment with antibiotics from one of three regional subcenters, or referral to higher-level medical treatment centers. Cases were identified both actively and passively by specially trained community health workers using a modified WHO case definition based on respiratory rates and other visible symptoms such as chest retractions (Stewart et al., 1994; WHO, 1991). The ALRI control program was establish within a geographically defined treatment area of Matlab. The remaining half of the population in the study area received standard government services and acts as a control population for this study.

Arsenic Exposure

The widespread contamination of drinking water sources by inorganic arsenic is the unintended consequence of a successful program begun in the 1970s that installed wells, also called tubewells, across the region in order to provide clean drinking water and prevent diarrheal diseases (M. Black, 1990; Smith et al., 2000). Arsenic contaminated water has no distinguishing color, smell, or taste, and it was not routinely tested for in well installations, so the contamination was not detected until health problems were identified. The magnitude of the problem was not fully realized until the mid-1990s (Smith et al., 2000). A survey of wells within Matlab conducted in 2002-2003 found 62 percent of over 13,000 wells had arsenic levels above the

Bangladeshi-government recommended level of 50 µg/L (ICDDRB, 2004). This contamination has been found to increase risks of skin lesions, hypertension, diabetes mellitus and is resulting in excess adult and infant mortality within Matlab (Yunus et al., 2011).

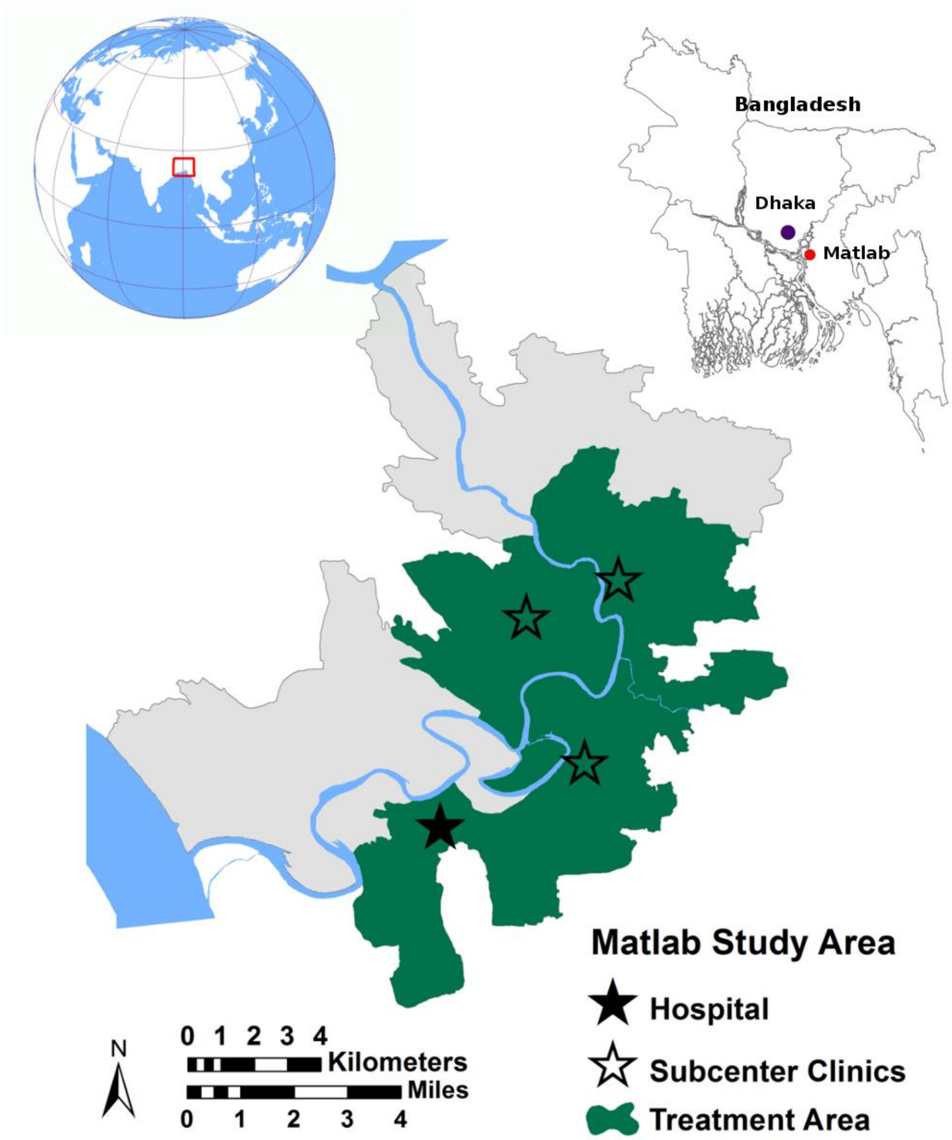

Figure 6: They study area of Matlab, Bangladesh showing the locations of the ALRI treatment area and medical facilities.

Study Data

Data for this study come from the demographic surveillance system (DSS) and the Matlab Geographic Information System (MGIS; Emch, 1999). The outcome is the mortality rate in children under 2 years old due to ALRI between 1989 and 1996. This period covers the years immediately following the implementation of the ALRI control program and before knowledge of arsenic contamination was widespread (Paul, 2004; Paul & De, 2000). To accurately estimate the mortality rate for children under 2 years, person-years of exposure were estimated using the DSS to identify all children and births that occurred within the study period. All children under 2 were followed in the DSS data across study years and internal migrations between baris until either death, out migration, or their second birthday occurred. Person-level mortality and exposure data were aggregated to the bari level, making the unit of analysis for this paper the bari. The original MGIS dataset contained 7846 baris, 1157 baris were excluded from the analysis because they either did not exist in the study area or had no children under 2 years during the study years, leaving a final dataset containing 6689 baris.

A bari-level estimate of arsenic exposure was created using data from the Matlab arsenic survey from 2002-2003 which measured arsenic by lab-based hydride generation atomic absorption spectrometry (HG-AAS) and located wells using global position system receivers (M. Rahman et al., 2006). Baris were assumed to experience the average concentration of arsenic in all wells owned by the bari. Following Carrel et al. (2011), well ownership was based on identification numbers linking wells to baris. If a bari did not have a privately owned well, the arsenic concentration of the nearest well was used. Nearest wells were identified in the MGIS using Euclidean distance between point locations of wells and baris. The census data records a household's primary source of drinking water (well, pond, canal, etc.). Using the 1996 census, any bari which did not have a household reporting well use was assumed to consume drinking water from a surface water source and was therefore assigned an arsenic exposure level of zero.

Socioeconomic status (SES) is closely tied to nutritional status in Matlab which can subsequently affect immune status (Stewart et al., 1994). SES was estimated using a principal components analysis (PCA) of a 1996 survey indicating the presence of several household assets (lamp, watch, radio, quilt) as well as house structure (walls/roof made of tin; use a latrine). The SES score for a bari was the average of all households' asset scores divided into quintiles. This procedure has been previously applied to demographic and health surveys in Matlab and elsewhere to estimate socioeconomic status in rural communities (Emch, Yunus, Escamilla, Feldacker, & Ali, 2010; Kolenikov & Angeles, 2009). For the analysis, socioeconomic status was dichotomized with the top two quintiles considered to be high status baris.

Using the MGIS, bari-level measures of access to care and population density were constructed. Cost distance to the nearest clinic or hospital is a measure of effort or 'cost' that it takes to travel a distance given the terrain. This measure is considered to reflect access to care within Matlab. The minimum cost distance was estimated using the procedure from Ali et al. (2001), which assumes that traveling a given distance across water requires five times the effort compared to land. Higher densities of people can increase the transmission of pathogens that cause ALRI (Cohen, 1999). The population density of the area surrounding each bari was calculated using a circular neighborhood with a radius of 200 meters centered on each bari. The all-age population of each bari in the 1993 census was used for the density calculation. The continuous variables, arsenic concentration, cost distance to clinic, and population density were centered and scaled before being entered into the analyses.

Statistical Analysis

Zero-inflated negative binomial (ZINB) regression was used to model the bari-level mortality rates. ZINB models are two-part mixture models that adjust for overdispersion (when the variance exceeds the mean) and excess zeros produced by rare outcomes (Cameron & Trivedi, 1998; Loeys et al., 2012). The log of the person-years of children under 2 was included as an offset population to account for differences in the population at risk across baris. All analyses were conducted using R 2.12.1 (R Core Development Team, 2010). ZINB regression was carried out with the *zeroinfl* procedure in the *pscl* package (Jackman, 2011; Zeileis, Kleiber, & Jackman, 2008).

Results

Between 1989 and 1996, 816 deaths from ALRI were reported in children under 2 years. These deaths represent 84 percent of ALRI deaths across all ages in Matlab. Out of 6689 baris included in the analysis, only 691 (10.3%) reported at least 1 death, producing a larger proportion (89.7%) of baris without any deaths (Figure 7). The distribution of mortality events was also slightly overdispersed (variance / mean = 1.278).

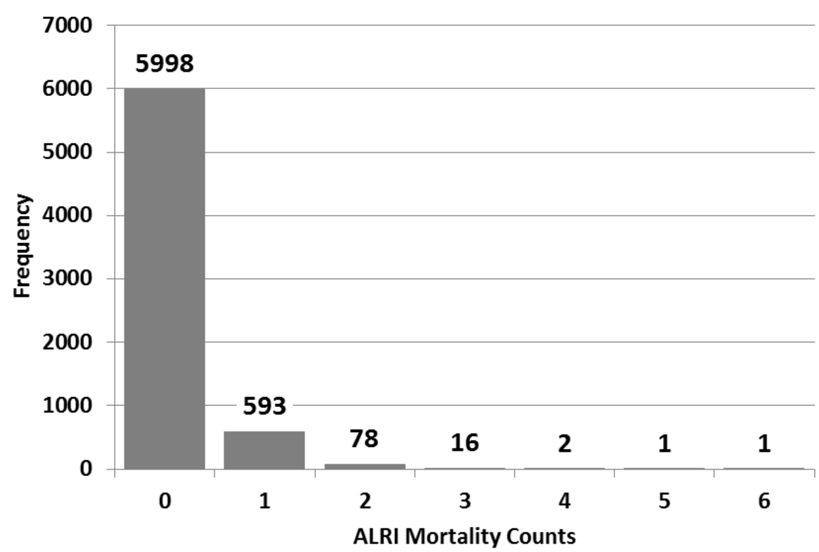

Figure 7: Distribution of the number of deaths in children under 2 years old due to Acute Lower Respiratory Infections (ALRI) reported by baris in Matlab, Bangladesh, 1989-1996.

Table 1 presents characteristics of the baris stratified by treatment and comparison area. ALRI mortality rates were almost 50 percent lower within the treatment area. While the treatment and comparison areas had similar average populations at risk in each bari, relative to the comparison area, the treatment area experienced lower average arsenic exposure levels, had a larger proportion (41%) of higher socioeconomic status baris, and a lower average population density. As all clinics are located within the treatment area, these baris also had lower average travel cost distances to reach a clinic. Use of tubewells and exposure to arsenic is widespread: only 130 baris (2%) reported using a water source other than a well in 1996 and 69.4% of all baris used in the study have estimated arsenic levels above the Bangladeshi-government recommended level of 50 μg/L.

Descriptive Comparisons for Baris in the Matlab Study Areas

	Treatment (n = 3,229)		Comparison (n = 3,460)		
ALRI mortality rate (per 1,000 PY)	6.15		10.67		
Total ALRI deaths 1989-1996	274		542		
Total person-years (children < 2)	44565.1		50798.5		
Total population (1993 Census)	103116		99534		
	mean	SD	mean	SD	p
Bari person-years (children < 2)	13.8	13.4	14.7	16.6	0.657
Average arsenic concentration (mcg/L)	185.9	181.3	252.4	225.0	0.000
High socioeconomic status	0.408	0.492	0.334	0.472	0.000
Cost distance to nearest clinic	2206.3	1161.2	5817	2232.6	0.000
Population density (population / sq. km)	2505.9	2149.2	2805.5	1633.2	0.000

Table 1: Bari characteristics of the treatment and comparison areas within Matlab, Bangladesh.

Table 2 shows the results of the ZINB analysis. The upper panel of the table presents coefficients and standard errors from a negative binomial model using a log link function of mortality in baris while the lower panel shows coefficients and standard errors from the zero-inflated logit model predicting that a bari reports not deaths. From the zero-inflated model, baris within the treatment area as well as baris further from a clinic were significantly more likely to report zero ALRI deaths. Baris with higher average concentrations of arsenic in their water were less likely to report no deaths, though this finding was only marginally significant ($p = 0.088$). Socioeconomic status of a bari was not a significant predictor of excess zeroes. After accounting for the excess of observed zeroes, the negative binomial model indicates that being located within the ALRI treatment area or being high socioeconomic status reduced mortality rates. Arsenic concentration and distance to clinic were not found to be significant predictors in the negative binomial model. The log(theta) term is a test of the additional shape parameter in a negative binomial model accounting for overdispersion. Population density was not found to be a significant predictor of mortality rates in earlier models and was not included in the final model presented here.

Negative Binomial Model

	B	SE	p	
Treatment Area	-0.579	0.108	0.000	***
Average Tubewell Arsenic	0.045	0.041	0.276	
High Socioeconomic Status	-0.282	0.090	0.002	**
Cost Distance to Clinic	-0.085	0.054	0.117	
Log(theta)	1.674	0.582	0.004	**
(Intercept)	-4.437	0.055	0.000	***

Zero-Inflated Model

	B	SE	p	
Treatment Area	10.467	4.095	0.011	*
Average Tubewell Arsenic	-4.106	2.409	0.088	+
High Socioeconomic Status	-3.679	4.315	0.394	
Cost Distance to Clinic	3.901	1.492	0.009	**
(Intercept)	-16.482	4.926	0.001	***

0.000 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

Table 2: Zero-inflated negative binomial regression analysis of bari-level acute lower respiratory infection (ALRI) mortality rates in children under 2 years of age in Matlab, Bangladesh, 1989-1996.

The results of the ZINB model can be expressed as the expected mean counts predicted by the negative binomial model. Figure 8 shows these counts predicted for varying levels of arsenic for a bari in the treatment and control areas while holding constant SES at low, and cost distance and person-years at their means. While the expected ALRI deaths increases with arsenic, at all levels of exposure baris in the treatment area, on average, experience lower numbers of ALRI deaths than the comparison area.

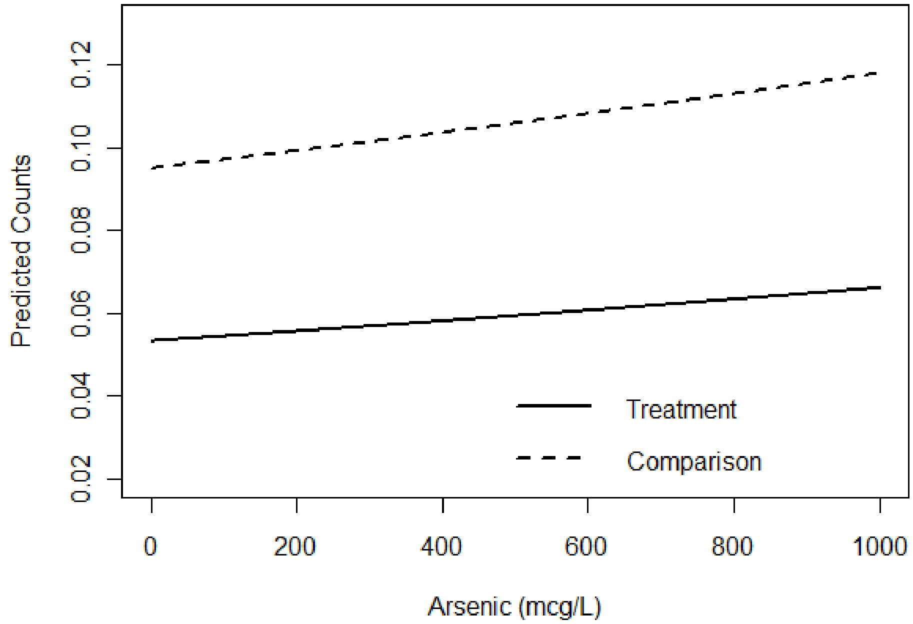

Figure 8: Predicted ALRI mortality in children under 2 years using the results of a negative binomial regression.

Discussion

While previous studies of acute lower respiratory infections have primarily focused on individual-level risk factors (e.g. De Francisco, Morris, Hall, Schellenberg, & Greenwood, 1993; Spika et al., 1989; Zaman et al., 1997), this research sought to highlight the broader contextual and environmental characteristics that can influence mortality rates. The objectives of this study were to evaluate these risk factors for childhood ALRI mortality in the contexts of a community-based control program and widespread exposure to inorganic arsenic from contaminated drinking water in Matlab, Bangladesh. This study found that living within the area served by the ALRI control program was strongly associated with reduced mortality rates measured at the bari level. This finding is consistent with previous studies (Ali et al., 2001; Fauveau et al., 1992; Stewart et al., 1994) that found that the control program in Matlab was highly successful at reducing mortality in young children by up to 50%. In the present study, we also hypothesized that exposure to arsenic would be associated with increased ALRI mortality. Studies in Matlab and elsewhere have found that arsenic exposure is associated with increases in adverse pregnancy outcomes and infant mortality (Vahter, 2009), as well as with harm to the lungs including coughs, cancer, and infections (Guha Mazumder, 2007; Raqib et al., 2009; Smith et al., 1998). Adding to the biological plausibility for the potential association with ALRI mortality are studies which have shown that arsenic's toxic effects suppress the immune system, particularly in children (Soto-Peña et al., 2006), and damages cellular DNA and chemical receptors in lung tissue providing the opportunity for infections (Kozul, Hampton, et al., 2009; Kozul, Ely, et al., 2009). However, in a zero-inflated negative binomial analysis used in this study, increased arsenic levels were only marginally associated ($p = 0.088$) with a reduced probability of not reporting an ALRI death. Higher mortality rates were found with increasing levels of arsenic, but this finding did not reach statistically significant levels ($p = 0.276$).

This study has extended a previous evaluation of the Matlab ALRI control program by Ali et al. (2001) by including three additional years of mortality records as well as by incorporating measures of socioeconomic status and arsenic exposure that have not been previously used. Baris with higher SES were associated with lower rates of child mortality from ALRI in this study after controlling fort the treatment area effect. An earlier case-control study on child ALRI mortality in Gambia did not find a significant association with SES (de Francisco et al., 1993), and, similarly, a prospective cohort study of children under 5 years in Matlab did not find an association with incidence of respiratory infections and various sociodemographic measures (Zaman et al., 1997). However, many studies have found various measures of social status including income, home ownership, and education to be predictive of respiratory infections in various age groups and country settings (Cohen, Doyle, Turner, Alper, & Skoner, 2004; Crighton et al., 2007; Margolis et al., 1992). The conflicting findings may be due to differences in the specific measure of SES used in each study. SES may affect respiratory infections by increasing exposure to pathogens in crowded living quarters or by decreasing an individual's immune status due to stress or poor nutrition (Cohen, 1999). The cost distance to the nearest treatment center was also found to be significantly associated with an increased probability of not reporting an ALRI death. Ali et al. (2001) similarly found that decreased access to care as measured by cost distance was associated with reduced mortality rates. While I expected that limiting access to care would increase mortality, this opposite finding is possible evidence of a reporting bias with more distant baris less likely to accurately report cases. Previous studies of diarrheal diseases in Matlab have found a similar pattern of decreased case numbers of diarrheal diseases with increasing distance from the hospital (Carrel et al., 2011, 2010). The study by Ali et al. (2001) also incorporated several measures of access to alternative allopathic and indigenous health care; however, these data were not available for this analysis.

Strengths of this study include the detailed records collected in the demographic surveillance system in Matlab which enables accurate reporting of ALRI deaths and estimates of the population at risk. That these records can be linked in a geographic information system to spatial data on housing locations and tubewells with measured arsenic concentrations further enhances this study and allows for additional environmental variables to be considered. A significant limitation of this study is the arsenic exposure measure. The arsenic data come from a more recent (2002 to 2003) survey of tubewells in Matlab which are applied to the earlier study period of 1989 to 1996. Evidence from several studies in Bangladesh, though limited, suggests that arsenic levels in wells are generally consistent over time (BGS & DPHE, 2001; Cheng, van Geen, Seddique, & Ahmed, 2005; Cheng et al., 2006); however, it is not known which specific wells existed and were used during the study period. As arsenic contamination became more well-known in the 1990s, wells installed between 1996 and the arsenic survey in 2002 were likely deeper to access clean water while older wells which may have been broken or removed by the survey year were typically shallower and, thus, more likely to be contaminated. These changes in wells over the years between the study period and arsenic survey would likely bias the arsenic exposure downward and averaging all tubewells belonging to a bari also potentially reduces the estimated exposure measure. These steps likely produce a conservative estimate of a bari's true arsenic exposure and may explain the lack of a significant association between arsenic exposure and ALRI mortality in this study.

As acute lower respiratory infections continue to be a major cause of illness and death for children in Bangladesh and around the world, it is important to evaluate the effectiveness of community-based intervention strategies on population health. This research contributes to our understanding of these programs by testing an additional potential risk factor of arsenic exposure and also by demonstrating the need to consider aggregate, contextual factors such as measures of socioeconomic status.

Bridge

The results of the zero-inflated negative binomial model presented in the first paper suggest that the ALRI control program had a strong effect reducing child mortality within the treatment area. After controlling for the treatment area effect, higher socioeconomic status was also associated with lower mortality rates. A bari's average arsenic exposure was only marginally associated with mortality. While these first analyses are able to investigate these risk factors, the models are not explicitly spatial. The ALRI program was implemented in a specific geographic area and likewise the distribution of the population and characteristics is not uniform across Matlab. Therefore there are possibly spatial differences in the mortality risk experienced by people across Matlab. Ali et al. (2001) mapped a smoothed surface of mortality rates in Matlab and found areas of locally higher rates. However, they did not examine whether higher rates could be due to chance or because of the distribution of various social and environmental characteristics. Identifying areas of locally higher or lower mortality rates and the factors associated with those spatial patterns can help to highlight disparities and priority areas for public health interventions. Moreover, these types of spatial studies are also useful for generating new hypotheses and future studies exploring the spatial differences. In the second paper I apply a spatial scan statistic (Kulldorff, 1997) to identify such areas of unexpected and unexplained spatial grouping of ALRI deaths after adjusting for varying population and covariate distributions across the study area. By using and interpreting the results of the first paper in conjunction with the second paper, I am able to examine the ALRI control program and address my research questions by identifying characteristics associated with ALRI and the spatial patterns of this mortality in the population.

Paper 2: Spatial Clustering of Acute Lower Respiratory Infections in Bangladeshi Children Using a Regression-adjusted Scan Statistic

For submission to *Health & Place*

Abstract

Respiratory infections continue to be a public health threat, particularly to young children in developing countries. A spatial scan statistic approach was used to assess the location and magnitude of local clusters of mortality from acute lower respiratory infections (ALRI) in young children (less than 2 years) in an area of rural Bangladesh. This method was combined with a regression model predicting mortality events in order to adjust for multiple social and environmental risk factors. Adjusting for access to specialized ALRI treatment explained the existence of a high and low cluster of mortality; yet further adjustments could not completely explain an area of lower than expected mortality.

Introduction

Pneumonia and other acute lower respiratory infections (ALRI) claim the lives of over 1.5 million children around the world each year (R. E. Black et al., 2010). However the burden of these diseases varies regionally: South and Southeast Asia suffer some of the highest mortality rates, with approximately 21 percent of all deaths in children under 5 years old attributed to pneumonia (R. E. Black et al., 2010). These macro-scale patterns are well-documented and do indicate disparities and priority areas needing medical and public health interventions; however, mortality rates can also vary significantly at more local scales. Prior studies have found that childhood ALRI are associated with poverty, malnutrition, indoor air pollution, crowded living conditions, as well as access to medical care (Ali et al., 2001; de Francisco et al., 1993; Lanata & Black, 2008; Rudan et al., 2008). These factors likely operate by affecting immune status or by increasing exposure to pathogens and lung irritants (Mizgerd, 2006; Rudan et al., 2008). The unique context of a place and the uneven distribution of these risk factors in a population can lead to local-scale spatial variation in ALRI.

In Bangladesh ALRI are a leading cause of morbidity and mortality among children (Ali et al., 2001; ICDDRB, 2010; Spika et al., 1989). Children on average suffer 0.23 and 0.47 events per year (Arifeen et al., 2009; Zaman et al., 1997) and over 25,000 die from pneumonia alone each year (R. E. Black et al., 2010). ALRI continues to be a health threat, though it is largely treatable and preventable. Vaccines for the two main causes of ALRI, pneumococcus and *Haemophilus influenzae* type b (Hib), can reduce mortality but they are not widely available to the most vulnerable populations due to cost and distribution challenges (Girard et al., 2005). In their absence, community-based health intervention programs designed to identify new cases of respiratory infection early and provide antibiotic treatments have been found to be effective at reducing child mortality from ALRI, especially in rural and developing areas where vaccine distribution may be limited (Sazawal & Black, 2003). One such program established in an area of Matlab, Bangladesh was found to reduce mortality from ALRI in children by up to 50 percent (Ali et al., 2001; Fauveau et al., 1992). Ali et al. (2001) also identified considerable geographic variation in the mortality rates experienced by groups of households in that community; however, their work was largely a descriptive visualization of rates to accompany a non-spatial analysis and they stopped short of testing whether those patterns could occur due to chance and what social and environmental factors which could generate the observed differences.

The objective of this study is to examine the effect of the ALRI control program on local child mortality patterns in Matlab, Bangladesh. Studying local disease patterns is vital to understanding the role of known risk factors as well as to explore possible additional factors generating spatial disparities in mortality in order to provide more targeted and appropriate public health interventions. Using a spatial scan statistic (Kulldorff, 1997) I test for the existence and location of statistically significant local clusters of mortality events. As the ALRI program was only implemented in one area of the region and was found to be successful in reducing mortality, it will necessarily impact the spatial patterns of ALRI. I hypothesize, however, that after controlling for the treatment area effect, additional conditions related to socioeconomic status and access to care will result in spatial clusters of childhood mortality. Any apparent clustering may be due to the uneven distribution of population and risk factors in the study area. Therefore the scan statistic is iteratively adjusted by adding risk factors and observing changes in the likelihood and location of any detectable clusters, allowing us to better understand the relationship between risk factors and ALRI. A secondary objective of this study is to demonstrate the ability to combine cluster detection tests with regression analyses to more accurately adjust the spatial scan statistic.

Background

The study area of Matlab is located approximately 50 km southeast of Dhaka in central Bangladesh (Figure 9). The rural region is home to a population of over 200,000 people who live in clusters of related households called *baris*. Matlab has been the site of an ongoing health and demographic surveillance system (HDSS) since the 1960s, operated by the International Center for Diarrheal Disease Research, Bangladesh (ICDDR,B). The HDSS records all vital events, including births, deaths, and migrations, for each individual in the study area. Periodic censuses and surveys also record socioeconomic data.

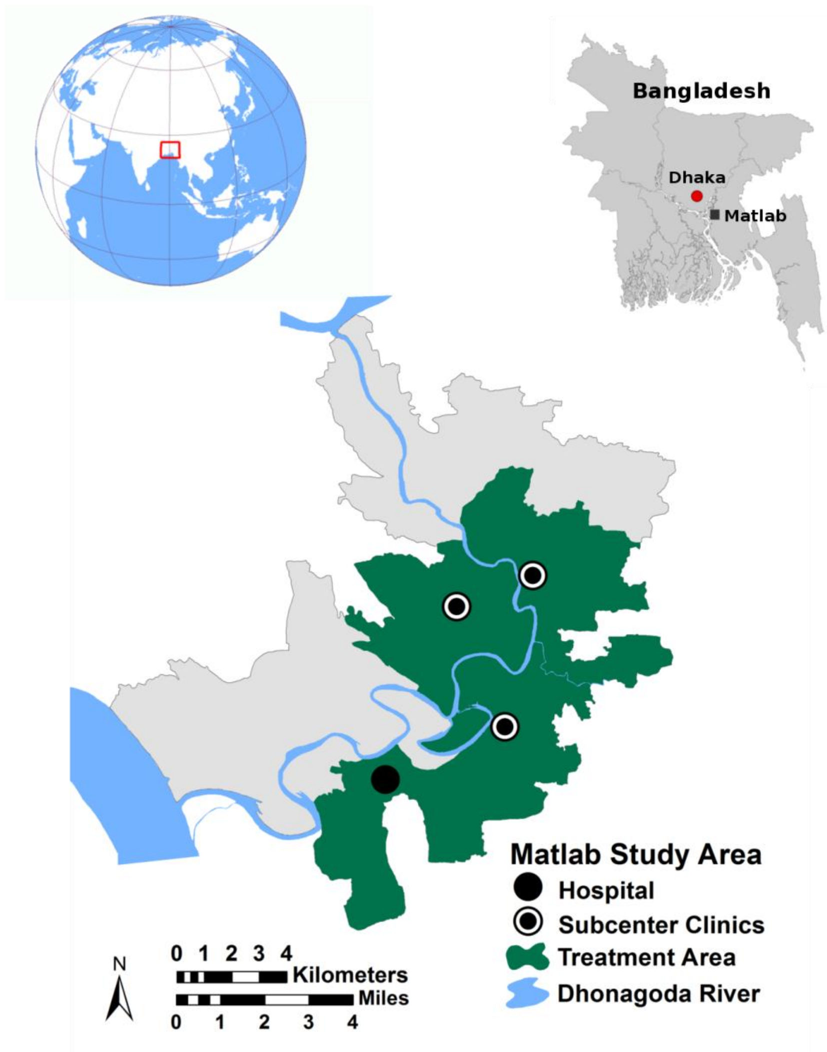

Figure 9: Matlab study area.

Starting in 1977 ICDDR,B initiated a series of interventions providing family planning services, immunizations, and perinatal care. Known as the maternal and child health and family planning program (MCH/FP), these interventions were successful in reducing fertility and mortality by increasing the prevalence of contraceptives and vaccine use (Fauveau, Wojtyniak, Chakraborty, Sarder, & Briend, 1990; Koenig, Bishai, & Khan, 2001; Koenig, Rob, Khan, Chakraborty, & Fauveau, 1992). The MCH/FP was implemented in a geographically defined area of Matlab (Figure 9). Approximately one half of the villages in Matlab are within the "treatment area" and received the specialized MCH/FP interventions. The other half of the population forms two comparison areas adjacent to the treatment area, one in the north and the other in the southwest, and received standard government services. Both the treatment and comparison area populations are recorded in the HDSS. Prior to the start of the MCH/FP the treatment and comparison areas had similar health and demographic measures. Beginning in 1989 the MCH/FP expanded to start an ALRI control program within the treatment area (Fauveau et al., 1992; Stewart et al., 1994). This community-based program was designed to reduce childhood mortality from ALRI through a combination of health education activities, early case detection, and management of cases using antibiotic treatments. Three outpatient, subcenter clinics as well as hospital facilities were created to treat ALRI (shown in Figure 9). The ALRI program was managed by trained community health workers (CHW) who regularly visited households to give caretakers information on symptoms and treatment of ALRI and to identify ALRI cases using a modified WHO case definition based on respiratory rates and other visible symptoms such as chest retractions (WHO, 1991; (Stewart et al., 1994). ALRI was classified as mild, moderate, or severe. Mild cases were monitored and mothers were given additional health information for supportive care. A moderate case, diagnosed as respiratory rates greater than 50 breathes per minute but without other symptoms, was treated with antibiotics. Severe pneumonia cases, defined by the presence of chest retractions and other symptoms along with respiratory rates above 50 breathes per minute or any lung infection in children under 1 month old regardless of symptoms, were referred to the ICDDR,B hospital in Matlab where oxygen, intravenous fluids and antibiotics were available.

Data for this study come from the mortality records of the HDSS and include all deaths reported from pneumonia or other ALRI. Similar to Ali et al. (2001), this study focuses on the population which experiences the most ALRI mortality, children under 2 years old, and during the years immediately following the start of the ALRI control program (1989-1996) to evaluate its effects. HDSS data for each individual in Matlab can be linked across study years to incorporate census and survey data as well as linked spatially to the bari locations in the Matlab Geographic Information System (MGIS; Emch, 1999). The person-years of children under age 2 years in each bari define the population at risk. Person-years are calculated using HDSS data by first identifying all births in the study area and immigrations of children under 2 years to Matlab. These individual children are also linked to their baris by their household identification numbers and then tracked through the migration and death files until one of three outcomes: they permanently out-migrate, die, or turn 2 years old. The time spent by all children under age 2 years for all study years is aggregated by bari. Deaths are also aggregated across study years and, along with the person-years of population at risk, are linked to baris. Baris are represented as point locations in the MGIS and become the unit of analysis for the spatial scan statistic. A total of 7846 baris were identified during the 1989-1996 study period; though 1157 of which did not contain any children and were excluded from the analysis leaving 6689 baris.

Additional covariates used when adjusting the scan statistic include the geographic placement of the treatment area, bari-level socioeconomic status, cost distance to an ALRI treatment center, and exposure to inorganic arsenic from contaminated drinking water. Area-level socioeconomic status (SES) has been shown to influence lower respiratory infections in other contexts (Cohen et al., 2004; Crighton et al., 2007), and, within Matlab, SES is linked with nutritional status which has implications for a child's immune system health (Stewart et al., 1994). Similar to previous studies (Alam, Zahirul Haq, & Streatfield, 2010; Emch et al., 2010; Kolenikov & Angeles, 2009) SES was measured using a principal components analysis of the presence of several household assets (lamp, watch, radio, quilt) as well as house structure (walls/roof made of tin; use a latrine) as recorded in a 1996 household survey. The bari-level SES score is the average of the first principal component score for all households in the bari. The bari SES is then divided into quintiles, and dichotomized with the top two quintiles considered high socioeconomic status.

Cost distance is a measure of accessibility or effort to reach a treatment center that considers both distance and physical barriers to reaching care. A similar measure was found to be significant in the earlier study of ALRI mortality by Ali et al. (2001). The cost distance to the nearest ALRI treatment center was calculated as a raster surface (30 m resolution) in the MGIS using the point locations of treatment subcenters or hospital and the boundaries of rivers and bodies of water in the study area. The cost distance assumes that travelling across water requires five times the effort of land-based travel.

A bari-level average arsenic exposure measure was also considered. Widespread contamination of water sources by inorganic arsenic is a serious health threat in Bangladesh that has been linked to infant mortality and increases in infectious diseases including respiratory infections (A. Rahman et al., 2007, 2010; Raqib et al., 2009; Yunus et al., 2011). Arsenic exposure was estimated using data from the Matlab arsenic survey conducted in 2002-2003 which measured arsenic and located wells using global position system receivers (M. Rahman et al., 2006). Baris were assumed to experience the average concentration of arsenic in all wells owned by the bari based on well identification numbers linking them to baris. If a bari did not have a privately owned well, the arsenic concentration of the nearest well was used. Nearest wells were identified in the MGIS using Euclidean distance measures between point locations of wells and baris. Using the 1996 socioeconomic census, any bari which did not report having a household using a well was assumed to have an arsenic level of zero.

Statistical Methodology

Cluster detection tests were performed using the spatial scan statistic implemented in the SaTScan software package (Kulldorff, 1997; Kulldorff & Information Management Services Inc., 2009). The spatial scan statistic operates by placing a large number of circles of varying radii at each location, and calculates the ratio of observed events to expected events in the population within each scanning window. A likelihood ratio test is calculated for each circle to test whether the observed to expected ratio within a scanning window is different from the risk in the population outside of the scanning window. The maximum likelihood ratio identifies the most likely cluster at a location and statistical significance is determined using Monte Carlo simulations of the locations of observed cases under the null hypothesis that events are distributed over the study area proportionally to the population. This study presents the results of three tests identifying purely spatial clusters of either high or low rates of ALRI mortality centered at bari point locations. I assumed the number of deaths to children under 2 years in each bari is Poisson distributed. The maximum scanning windows was limited to up to 50 percent of the population at risk. Clusters were considered to be significantly different from the null hypothesis of complete spatial randomness at the alpha = 0.1 level as determined by 999 Monte Carlo simulations. Only clusters with no geographic overlap are presented.

In all three analyses in this study – one unadjusted and two adjusted models – the observed number of deaths in a bari remains constant and is derived from the HDSS records (as described above). The expected number of deaths at a location, used as the denominator is the spatial scan calculation, varies as covariate adjustments are made. These changes in the expected cases and subsequent changes in the risk within a scanning window are the basis for interpreting the relationships between risk factors and disease patterns. If a covariate is positively related to an increase in mortality rate, the expected number of deaths will be increased and the observed to expected ratio will be reduced compared with a non-adjusted analysis. Thus, if a significant high cluster found in a given location in an unadjusted analysis is no longer significant after introduction of the covariate adjustment, we can say that the observed cluster was due to the uneven spatial distribution of that risk factor. Clusters which persist or appear after adjustment are not fully explained by a given model. Therefore cluster detection can be useful for generating new research questions and hypotheses regarding additional risk factors for a given disease process. The discrete Poisson model implemented in SaTScan allows for adjustment using only categorical covariates (Kulldorff, 1997, 2010). Within the software, the adjustment procedure calculates the expected number of events for a location in each covariate level using indirect standardization based on the supplied population or person-years at risk; however, this procedure can be too limiting for more complex analyses. When many covariates are entered into a model, particularly when multiple categories have no population or cases, the estimates may be unstable (Kulldorff, 2010). Furthermore, continuous covariates must be cut into strata before being used. Ideally these levels are determined *a priori* from a theoretical basis, but this step may introduce bias or conceal a significant effect. This study differs from most previous applications of the spatial scan statistic by adding a preprocessing step to calculate expected mortality counts using a separate regression analysis. While the observed counts remain fixed, by using separate regression software, the expected counts can be estimated from a more flexible and complete regression equation modeling the disease process and then supplied to the spatial scan statistic instead of the population at risk. These regressions have the potential to use continuous covariates as well as additional terms such as polynomials or interactions, or to be based on more sophisticated models such as multilevel models that incorporate individual and area level factors (Klassen et al., 2005).

The first analysis is unadjusted to provide a baseline for comparison. Without adjustment, the expected number of deaths in a bari is proportional to the population of children at risk. As the ALRI control program was implemented in a geographically defined area of Matlab and did reduce mortality, we can be confident that it will also affect spatial mortality patterns. The second analysis introduces an adjustment for whether a bari is located within the treatment area. In this analysis the expected mortality events are calculated using the treatment/comparison area-specific rates. The third analysis uses expected counts calculated from a previously constructed zero-inflated negative binomial regression that adjusts for treatment area effects, as well as bari-level socioeconomic status, cost distance to a medical facility, and exposure to inorganic arsenic from contaminated drinking water (see Paper 1). Using this model I am able to incorporate continues measures of the cost surface and arsenic exposure, as well as the categorical variables for treatment area and high socioeconomic status. Results from all three analyses are presented graphically after importing the results from SaTScan into ArcGIS, to identify the locations of significant clusters as well as in tabular format to compare changes in likelihood and relative risk among clusters.

Results

Within the 6689 baris retained for the analysis, a total of 816 deaths from ALRI were reported in children under 2 years old from 1989 to 1996. ALRI deaths were reported in 691 baris (10.3%), and the number of deaths in baris reporting an ALRI death ranged from 1 to 6.

Figure 10 and related Table 3 show the results of the unadjusted scan statistic which identified two statistically significant clusters. The most likely cluster is a large area of lower risk (relative risk, RR = 0.55, p = 0.000) centered over the treatment area. The second most likely cluster is an area of elevated risk (RR = 1.47, p = 0.027) in the southwest area of Matlab in the comparison area. Figure 11 and Table 3 show the results after adjusting for the treatment area placement. The most likely cluster is now an area of lower risk (RR = 0.39, p = 0.003) in the southwestern edge of the study area. There are no longer any areas of significantly elevated risk and the two clusters initially identified in the unadjusted analysis are no longer significant. The final model uses the full regression model to adjust for treatment area, socioeconomic status, cost distance, and arsenic exposure. Figure 12 and Table 3 indicate that only minor changes occurred after this third adjustment step. Similar to the second model, an area of lower risk (RR = 0.38, p = 0.090) is again found in the southwestern area of Matlab; however, the position of the cluster is shifted slightly and this cluster is smaller in diameter and contains fewer observed and expected cases than the treatment area adjusted cluster. This cluster is also borderline significant (p = 0.090) indicating that adjustment covariates did explain some of the decrease in risk found in this area.

	Radius (km)	Observed Deaths	Expected Deaths	Relative Risk	Likelihood Ratio	p-Value
Unadjusted						
(Figure 10)						
Cluster 1	4.89	162	254.11	0.548	26.35	0.000
Cluster 2	3.54	260	197.32	1.466	12.33	0.027
Treatment Area Adjusted						
(Figure 11)						
Cluster 1	1.41	25	60.95	0.392	14.51	0.003
Model Adjusted						
(Figure 12)						
Cluster 1	1.33	18	45.55	0.382	11.33	0.090

Table 3: Spatial cluster analysis of childhood deaths from acute lower respiratory infections (ALRI) in Matlab, Bangladesh, 1989-1996.

Discussion

The initial unadjusted spatial scan statistic identified two areas of significant spatial clustering – one each of higher and lower than expected risk of ALRI mortality. The high risk cluster has a 47 percent excess compared to the rest of the study area while the low risk cluster sees a reduction of 45 percent, which is similar to the 50 percent reduction reported by Ali et al. (2001) for the period from 1989 to 1993. The locations of these clusters visually appeared to be associated with the ALRI treatment area with lower risk in the central treatment area and higher risk in the southern comparison area. The earlier study by Ali et al. (2001) mapped mortality rates and found fewer areas of elevated rates in the treatment area. The present study also detected that pattern and the spatial scan statistic also allows for statistical inferences which show that the spatial clustering is significant and that controlling for the treatment area explains this variation.

Previous studies of mortality in Matlab using the spatial scan statistic have found similar patterns and areas of clustered mortality. In a village-level analysis of all-cause mortality in children under 5 years between 1998 and 2002, Alam et al. (2010) found significant clusters of elevated mortality centered on the northern and southern areas of Matlab as well as a secondary cluster of high risk along the eastern edge of the study area, after adjusting a space-time scan statistic for education and economic status. Using a spatial scan statistic on fetal and infant deaths between 1991 and 2000 and adjusting for age, parity, education and SES, Sohel et al. (2010) found a large cluster of significantly lower rates in central Matlab, and a smaller, but also significant cluster of elevated rates, in the southwestern portion of the study area. Visually these results are remarkably similar to those found in the unadjusted analysis in this study. Sohel and colleagues suggest that the clustering in fetal and infant mortality could be due to differences in arsenic exposure as they found significantly higher levels of arsenic in the wells used within the higher risk cluster. However, they did not adjust the scan statistic for arsenic level and test whether they could explain their observed spatial variation. In the present study, after adjusting for whether a bari was within the treatment area to receive ALRI interventions the clusters were no longer significant. Arsenic concentration varies significantly between the treatment and comparison areas ($\mu_{treatment} = 185.9$, SD = 181.3 vs. $\mu_{comparison} = 252.4$, SD = 225.0, $p < 0.000$; see Paper 1), and a measure of arsenic exposure was incorporated into the final model-adjusted scan statistic; however, in a previous study (see Paper 1) it was only found be marginally associated with ALRI mortality after adjusting for the differences in the population at risk, treatment area, SES, and cost distance. Therefore, I do not expect that arsenic alone would explain the clustering observed in the present study. The geographic pattern of the ALRI program implementation appears to be crucial for understanding the pattern of mortality in Matlab

As a sensitivity test, the unadjusted model was re-run using the option to only report clusters that contain less-than or equal-to 10% of the population at risk rather that the initial setting of 50% of the population at risk. Since the initial run found large clusters covering much of the study area, I was concerned they could be concealing smaller, still statistically significant, but less likely clusters, given the dominating pattern of the larger clusters and the conservation option of not allowing overlapping clusters. This additional test (Figure 13 and Table 4) did not find any additional clusters that did not fit the already reported pattern of an area of low risk in the treatment area and an area of higher risk in the southern comparison area. This additional test also suggests that the smaller, localized areas of high ALRI mortality rates mapped by Ali et al. (2001) are not likely to be significantly different than what would be expected due to chance.

After adjusting for the treatment area, which explained the initial clusters, a new cluster of lower ALRI mortality emerged in the southwestern edge of the study area. This area was only partially explained and the pattern persisted with only minor changes in size and position after the final scan statistic was adjusted for multiple social and environmental covariates. This cluster of lower risk was unexpected, and given its persistence after adjustment for covariates it represents and area of unexplained variation or model misspecification which will be explored in future research. The cluster could be the result of underreporting of ALRI deaths due to its location on the far edge of the study area. Previous studies of diarrheal disease have found fewer cases reported in the outlying areas of Matlab (Carrel et al., 2010). But given that these are mortality events it seems less likely that the death of a child would be missed. Another possibility is that these baris are receiving care elsewhere (e.g. outside of the Matlab study area) and so are not affected by the Matlab program placement. Ali et al. (2001) found that reduced ALRI mortality was associated with greater access to local allopathic doctors. Data on the locations of allopathic doctors were not available for this study, but if there is a greater availability or use of such care in this area compared with other areas of Matlab, it might explain the reduced risk.

The results of this study indicate that child mortality form ALRI does cluster spatially in Matlab. However these patterns are largely explained by the placement of the treatment program. These results confirm the findings from previous non-spatial studies that the ALRI control program is effective at reducing mortality and that, in the presence of such an effective control program, other socioeconomic and environmental factors are less influential to the mortality patterns. This work also provides one of the few examples of using regression models to incorporate continuous covariates and more complex models forms thereby improving adjustments to the spatial scan statistic. More broadly, this study demonstrates the importance of geographic studies to highlight areas of significantly elevated or reduced mortality in order to evaluate and improve public health intervention programs.

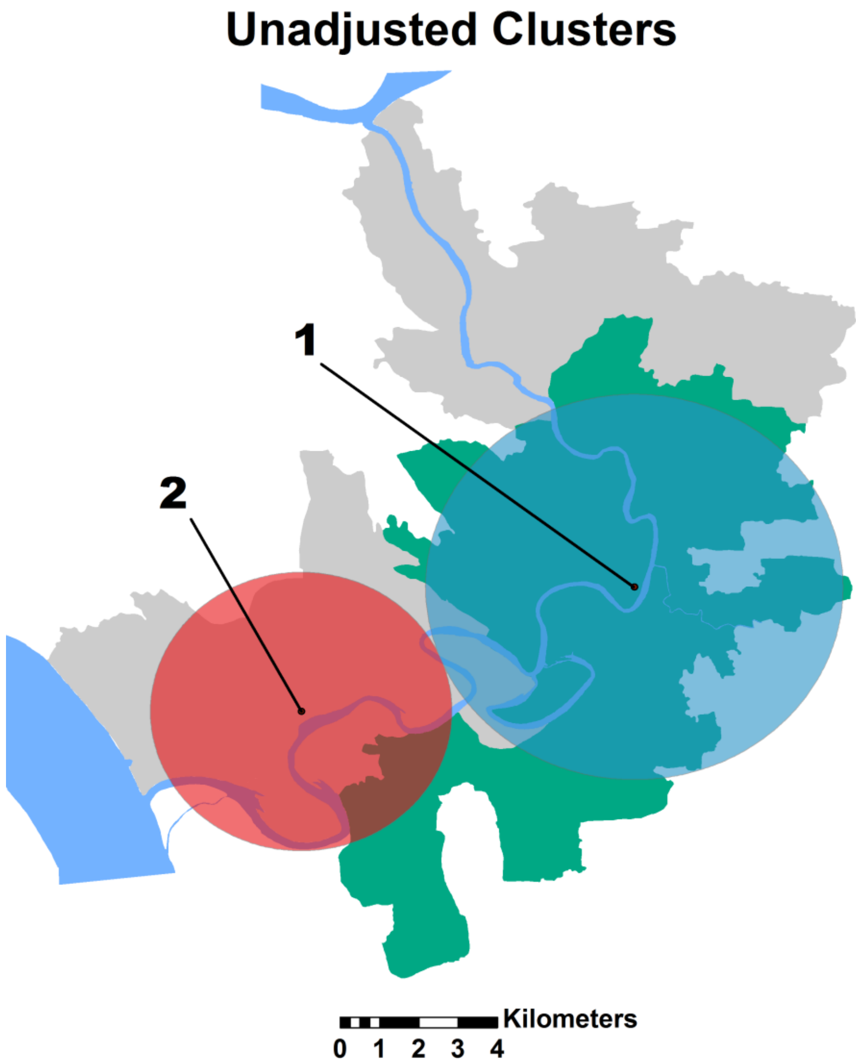

Figure 10: Results of unadjusted spatial scan statistic showing locations of statistically significant non-overlapping clusters of high (red) and low (blue) relative risk of mortality from acute lower respiratory infections.

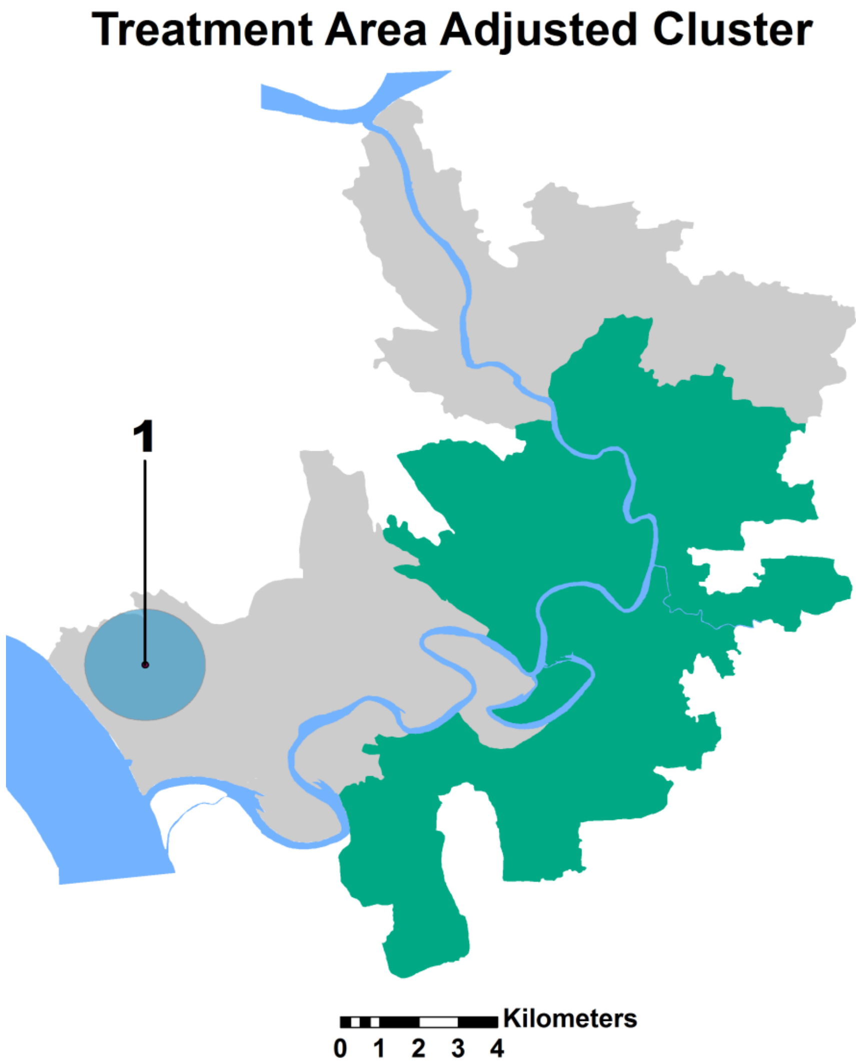

Figure 11: Results of a spatial scan statistic adjusted for the location of the treatment area (in green) showing locations of statistically significant non-overlapping clusters of low (blue) relative risk of mortality from acute lower respiratory infections.

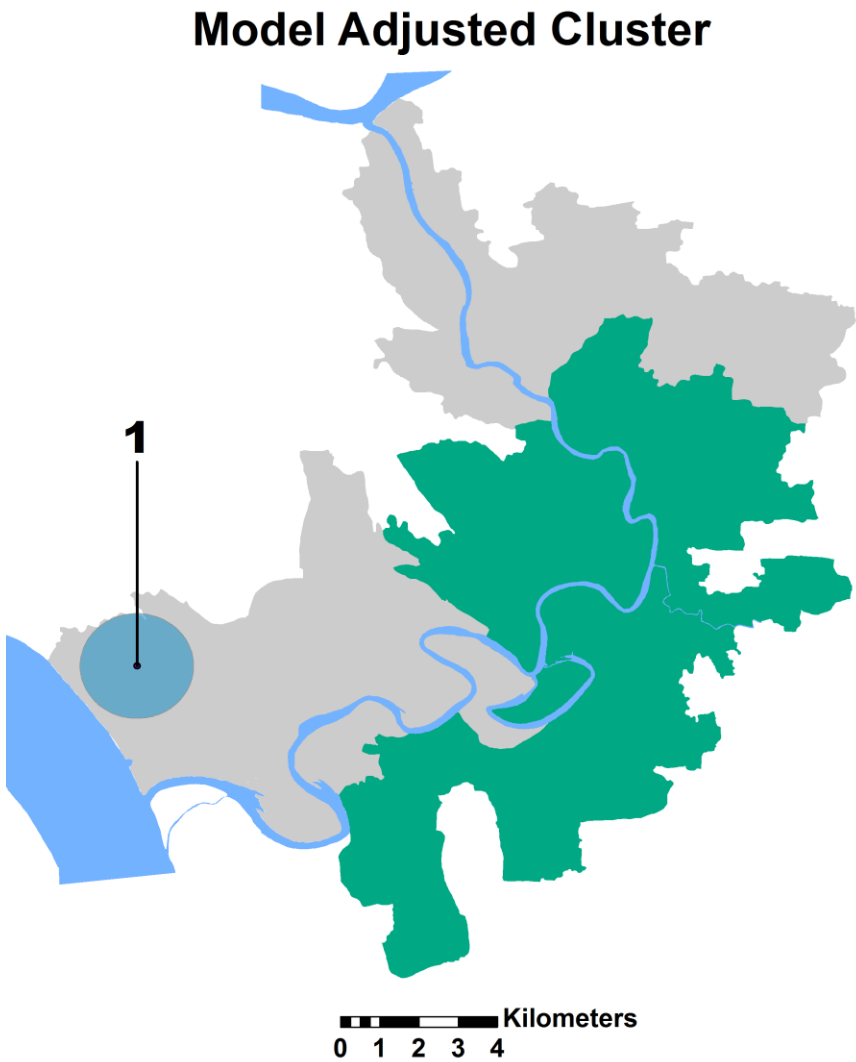

Figure 12: Results of a spatial scan statistic adjusted using a separate regression equation to calculate the expected events after adjusting for social and environmental factors. The location of one statistically significant cluster of lower relative risk is shown.

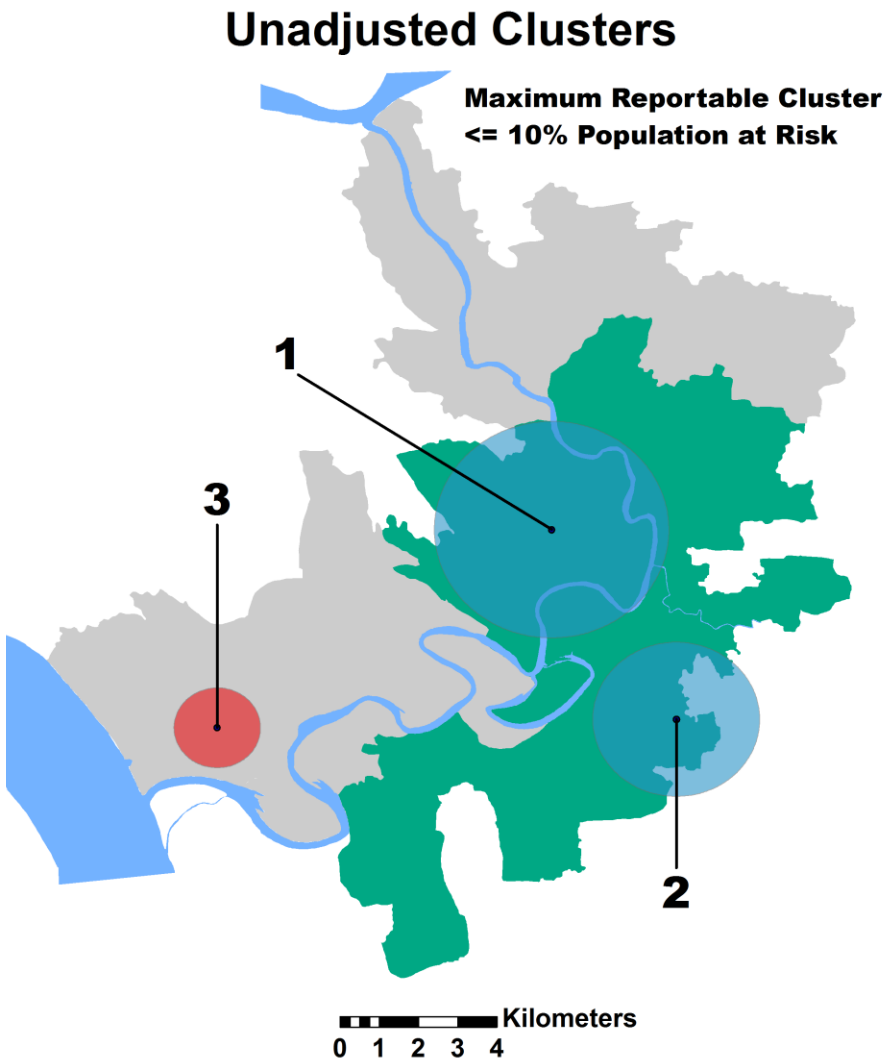

Figure 13: Results of an unadjusted spatial scan statistic limited to only reporting clusters contain less than 10% of the population at risk.

	Radius (km)	Observed Deaths	Expected Deaths	Relative Risk	Likelihood Ratio	p-Value
Unadjusted*						
(Figure 13)						
Cluster 1	2.75	36	77.49	0.440	15.04	0.002
Cluster 2	1.96	8	32.65	0.238	13.78	0.010
Cluster 3	1.02	64	33.45	1.991	11.58	0.055

* Maximum reportable cluster set to <= 10% of the population at risk

Table 4: Spatial cluster analysis of childhood deaths from acute lower respiratory infections (ALRI) in Matlab, Bangladesh, 1989-1996 after limiting reportable clusters to be <= 10% of the population at risk.

Conclusions and Discussion

In this study I have sought to clarify the effects of arsenic exposure and a community-based case management strategy on child mortality from acute lower respiratory infections in Matlab, Bangladesh. These analyses were focused on understanding the population-level characteristics and patterns of ALRI mortality in a disease ecology framework. This theoretical approach emphasizes the need to consider the combined interactions of biological, environmental and social/behavioral systems over space. In this work I have conceptualized the introduction of the ALRI control program and arsenic exposure from contaminated tubewells as two separate disturbances to the disease ecology of the region which have had positive and negative consequences for ALRI mortality. Because both arsenic and the treatment area exhibit variation and patterns at the bari level, I have tried to highlight the geographic variation of these exposures, and other social and environmental risk factors, to explore patterns childhood mortality.

I hypothesized that higher exposure to inorganic arsenic in contaminated drinking water would be associated with increased mortality from ALRI after controlling for the treatment area effect and that the spatial variation in arsenic exposure would contribute to the local-scale spatial patterns of mortality. These hypotheses were based on several studies from India, Bangladesh, and Chile where the toxic effects of arsenic have been linked to lung disorders and mortality in children and adults (von Ehrenstein et al., 2005; Guha Mazumder et al., 2000, 2005; Raqib et al., 2009; Smith et al., 2006). The relationship between child mortality from ALRI and arsenic has not been previously examined. In these two analyses I found that the treatment area was successful in reducing mortality and its geographic placement in Matlab played a dominant role in the spatial pattern of mortality. In addition, though, bari-level socioeconomic status measures were also associated with a reduction in mortality. This relationship to SES has not been documented in the previous studies (Ali et al., 2001; Fauveau et al., 1992; Stewart et al., 1994) of ALRI mortality in Matlab. Higher SES may be related to improved nutrition, improved or less-crowded living conditions, or health care seeking behaviors (Arifeen et al., 2008). The pathway through which a bari's SES level improves survival should be addressed in future studies. The second paper also identified an area of lower risk within the comparison area. This cluster could not be fully explained by a model incorporating socio-demographic or treatment and access to care, or arsenic exposure. This remaining cluster and the factors influencing lower mortality in this area could also be the subjects of future studies.

The results of the zero-inflated and clustering analyses do not support my original hypothesis that arsenic exposure will have a large effect on mortality patterns at the bari level. It is important to consider these results in the context of certain limitations in estimating arsenic exposure caused by baris changing wells over time and questions still remain about the ability to apply current arsenic measures to earlier exposure periods due to unknown long-term variation in arsenic levels. This study has focused on mortality patterns and cannot rule out a possible association between arsenic and ALRI morbidity. Moreover, future studies of individual-level exposure and ALRI outcomes would also be valuable, though as this study has shown, bari-level contextual factors such as SES should also be considered.

The analyses presented in these two papers also demonstrate two methodologies that are underutilized in health studies – zero-inflated models and a spatial scan statistic adjusted with a regression pre-processing step. Zero-augmented models are gaining wider use in health studies (Bohning, Dietz, Schlattmann, Mendonca, & Kirchner, 1999; Carrel et al., 2011; Rose, Martin, Wannemuehler, & Plikaytis, 2006) as they provide several advantages for handling excess zeroes along with over-dispersion commonly observed in counts and rates of rare events (Cameron & Trivedi, 1998). While the spatial scan statistic has been widely applied in health studies, there are only three instances in the literature in which a separate regression model has been used to adjust the scan statistic. The first (Huang et al., 2007) estimated and adjusted the survival times of an exponential model; the second (Kleinman, Abrams, Kulldorfff, & Platt, 2005), modeled temporal variation in the estimated population at risk; and the third (Klassen et al., 2005), used a multi-level model to estimate the expected cases of prostate cancer in census areas. Zero-inflated models have not been previously used with the spatial scan statistic. The regression-based adjustment method for the spatial scan statistic is easy to implement and has the potential to improve future studies by providing greater flexibility to describe risk factors.

Despite lessons learned from the experiences with the ALRI control program in Matlab, respiratory diseases continue to be a leading cause of death in Matlab (ICDDRB, 2010) and around the world (R. E. Black et al., 2010) and yet they receive relatively little attention (Mizgerd, 2006). The continued global burden of ALRI morbidity and mortality makes it important to evaluate intervention programs alongside social and environmental risk factors, such as arsenic exposure, in order to improve child survival.

References

Ahmed, S., Mahabbat-e Khoda, S., Rekha, R. S., Gardner, R. M., Ameer, S. S., Moore, S., Ekström, E.-C., et al. (2011). Arsenic-associated oxidative stress, inflammation, and immune disruption in human placenta and cord blood. *Environmental Health Perspectives, 119*(2), 258-264.

Alam, N., Zahirul Haq, M., & Streatfield, P. K. (2010). Spatio-temporal patterns of under-five mortality in Matlab HDSS in rural Bangladesh. *Global Health Action, 3*, 64-69.

Ali, M., Emch, M., Donnay, J. P., Yunus, M., & Sack, R. B. (2002). Identifying environmental risk factors for endemic cholera: A raster GIS approach. *Health & Place, 8*, 201-210.

Ali, M., Emch, M., Tofail, F., & Baqui, A. H. (2001). Implications of health care provision on acute lower respiratory infection mortality in Bangladeshi children. *Social Science & Medicine, 52*, 267-277.

Andrew, A. S., Bernardo, V., Warnke, L. A., Davey, J. C., Hampton, T., Mason, R. A., Thorpe, J. E., et al. (2007). Exposure to arsenic at levels found in U.S. drinking water modifies expression in the mouse lung. *Toxicological Sciences*, *100*(1), 75-87.

Andrew, A. S., Jewell, D. A., Mason, R. A., Whitfield, M. L., Moore, J. H., & Karagas, M. R. (2008). Drinking-water arsenic exposure modulates gene expression in human lymphocytes from a U.S. population. *Environmental Health Perspectives*, *116*(4), 524-31.

Argos, M., Kalra, T., Rathouz, P. J., Chen, Y., Pierce, B., Parvez, F., Islam, T., et al. (2010). Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): A prospective cohort study. *Lancet*, *376*(9737), 252-258.

Arifeen, S. E., Baqui, A. H., Victora, C. G., Black, R. E., Bryce, J., Hoque, D. M. E., Chowdhury, E. K., et al. (2008). Sex and socioeconomic differentials in child health in rural Bangladesh: Findings from a baseline survey for evaluating integrated management of childhood illness. *Journal of Health, Population, and Nutrition*, *26*(1), 22-35.

Arifeen, S. E., Saha, S. K., Rahman, S., Rahman, K. M., Rahman, S. M., Bari, S., Naheed, A., et al. (2009). Invasive pneumococcal disease among children in rural Bangladesh: Results from a population-based surveillance. *Clinical Infectious Diseases*, *48*(Suppl 2), S103-S113.

Atkins, P. J., Hassan, M. M., & Dunn, E. (2006). Toxic torts: Arsenic poisoning in Bangladesh and the legal geographies of responsibility. *Transactions of the Institute of British Geographers, 31*, 272-285.

BGS & DPHE. (2001). *Arsenic Contamination of Groundwater in Bangladesh*. (D. G. Kinniburgh & P. L. Smedley, Eds.) *British Geological Survey (BGS) Technical Report WC/00/19,* (2nd ed., Vol. 2). Keyworth: British Geological Survey.

Banerjee, N., Banerjee, S., Sen, R., Bandyopadhyay, A., Sarma, N., Majumder, P., Das, J. K., et al. (2009). Chronic arsenic exposure impairs macrophage functions in the exposed individuals. *Journal of Clinical Immunology*, *29*(5), 582-94.

Bearak, B. (1998, November 10). New Bangladesh disaster: Wells that pump poison. *The New York Times*. New York, NY.

Berg, M., Tran, H. C., Nguyen, T. C., Pham, H. V., Schertenleib, R., & Giger, W. (2001). Arsenic contamination of groundwater and drinking water in Vietnam: A human health threat. *Environmental Science & Technology*, *35*(13), 2621-2626.

Black, M. (1990). *From Handpumps to Health: The Evolution of Water and Sanitation Programmes in Bangladesh, India and Nigeria*. New York, NY: United Nations Children's Fund.

Black, R. E., Cousens, S., Johnson, H. L., Lawn, J. E., Rudan, I., Bassani, D. G., Jha, P., et al. (2010). Global, regional, and national causes of child mortality in 2008: A systematic analysis. *Lancet*, *375*(9730), 1969-1987.

Bloom, M. S., Fitzgerald, E. F., Kim, K., Neamtiu, I., & Gurzau, E. S. (2010). Spontaneous pregnancy loss in humans and exposure to arsenic in drinking water. *International Journal of Hygiene and Environmental Health*, *213*(6), 401-413.

Bohning, D., Dietz, E., Schlattmann, P., Mendonca, L., & Kirchner, U. (1999). The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. *Journal of the Royal Statistical Society: Series A*, *162*(2), 195-209.

Caldwell, B. K., Caldwell, J. C., Mitra, S. N., & Smith, W. (2003). Tubewells and arsenic in Bangladesh: Challenging a public health success story. *International Journal of Population Geography, 9*(1), 23-38.

Cameron, A. C., & Trivedi, P. K. (1998). *Regression Analysis of Count Data*. Cambridge, UK: Cambridge University Press.

Carrel, M., Escamilla, V., Messina, J., Giebultowicz, S., Winston, J., Yunus, M., Streatfield, P. K., et al. (2011). Diarrheal disease risk in rural Bangladesh decreases as tubewell density increases: A zero-inflated and geographically weighted analysis. *International Journal of Health Geographics*, *10*(41).

Carrel, M., Voss, P., Streatfield, P. K., & Yunus, M. (2010). Protection from annual flooding is correlated with increased cholera prevalence in Bangladesh: A zero-inflated regression analysis. *Environmental Health*, *9*(13), 1-9.

Cheng, Z., van Geen, A., Seddique, A. A., & Ahmed, K. M. (2005). Limited temporal variability of arsenic concentrations in 20 wells monitored for 3 years in Araihazar, Bangladesh. *Environmental Science & Technology*, *39*(13), 4759-4766.

Cheng, Z., van Geen, A., Seddique, A. A., & Ahmed, K. M. (2006). Response to comments on "Limited Temporal Variability of Arsenic Concentration in 20 Wells Monitored for 3 Years in Araihazar, Bangladesh." *Environmental Science & Technology*, *40*(5), 1718-1720.

Cohen, S. (1999). Social status and susceptibility to respiratory infections. *Annals of the New York Academy of Sciences*, *896*(1), 246–253.

Cohen, S., Doyle, W. J., Turner, R. B., Alper, C. M., & Skoner, D. P. (2004). Childhood socioeconomic status and host resistance to infectious illness in adulthood. *Psychosomatic Medicine*, *66*(4), 553-558.

Concha, G., Vogler, G., Lezcano, D., Nermell, B., & Vahter, M. (1998). Exposure to inorganic arsenic metabolites during early human development. *Toxicological Sciences*, *44*(2), 185-190.

Crighton, E. J., Elliott, S. J., Moineddin, R., Kanaroglou, P., & Upshur, R. (2007). A spatial analysis of the determinants of pneumonia and influenza hospitalizations in Ontario (1992-2001). *Social Science & Medicine*, *64*(8), 1636-50.

Dhar, R. K., Zheng, Y., Stute, M., van Geen, A., Cheng, Z., Shanewaz, M., Shamsudduha, M., et al. (2008). Temporal variability of groundwater chemistry in shallow and deep aquifers of Araihazar, Bangladesh. *Journal of Contaminant Hydrology*, *99*(1-4), 97-111.

Dubos, R. (1965). *Man Adapting*. New Haven, CT: Yale University Press.

von Ehrenstein, O. S., Mazumder, D. N. G., Yuan, Y., Samanta, S., Balmes, J., Sil, A., Ghosh, N., et al. (2005). Decrements in lung function related to arsenic in drinking water in West Bengal, India. *American Journal of Epidemiology*, *162*(6), 533-541.

Elliott, P., & Wakefield, J. (2001). Disease clusters: Should they be investigated, and, if so, when and how? *Journal of the Royal Statistical Society: Series A*, *164*(1), 3-12.

Emch, M. (1999). Diarrheal disease risk in Matlab , Bangladesh. *Social Science & Medicine*, *49*, 519-530.

Emch, M., Yunus, M., Escamilla, V., Feldacker, C., & Ali, M. (2010). Local population and regional environmental drivers of cholera in Bangladesh. *Environmental Health*, *9*(2), 1-8.

Fauveau, V., Stewart, M. K., Chakraborty, J., & Khan, S. A. (1992). Impact on mortality of a community-based programme to control acute lower respiratory tract infections. *Bulletin of the World Health Organization*, *70*(1), 109-116.

Fauveau, V., Wojtyniak, B., Chakraborty, J., Sarder, a M., & Briend, A. (1990). The effect of maternal and child health and family planning services on mortality: Is prevention enough? *BMJ*, *301*(6743), 103-107.

de Francisco, A., Morris, J., Hall, A. J., Schellenberg, J. R. M. A., & Greenwood, B. M. (1993). Risk factors for mortality from acute lower respiratory tract infections in young Gambian children. *International Journal of Epidemiology*, *22*(6), 1174-1182.

van Geen, A., Ahsan, H., Horneman, A. H., Dhar, R. K., Zheng, Y., Hussain, I., Ahmed, K. M., et al. (2002). Promotion of well-switching to mitigate the current arsenic crisis in Bangladesh. *Bulletin of the World Health Organization*, *80*(9), 732-737.

van Geen, A., Zheng, Y., Versteeg, R., Stute, M., Horneman, A., Dhar, R. K., Steckler, M., et al. (2003). Spatial variability of arsenic in 6000 tube wells in a 25 km2 area of Bangladesh. *Water Resources Research*, *39*(5), 1-16.

Girard, M. P., Cherian, T., Pervikov, Y., & Kieny, M. P. (2005). A review of vaccine research and development: Human acute respiratory infections. *Vaccine*, *23*(50), 5708-5724.

Glass, G. E. (2000). Spatial aspects of epidemiology: The interface with medical geography. *Epidemiologic Reviews*, *22*(1), 136-139.

Guha Mazumder, D. N. (2007). Arsenic and non-malignant lung disease. *Journal of Environmental Science and Health, Part A*, *42*(12), 1859-67.

Guha Mazumder, D. N. (2008). Chronic arsenic toxicity & human health. *Indian Journal of Medical Research*, *128*, 436-447.

Guha Mazumder, D. N., Haque, R., Ghosh, N., De, B. K., Santra, A., Chakraborti, D., & Smith, A. H. (2000). Arsenic in drinking water and the prevalence of respiratory effects in West Bengal, India. *International Journal of Epidemiology*, *29*, 1047-1052.

Guha Mazumder, D. N., Steinmaus, C., Bhattacharya, P., von Ehrenstein, O. S., Ghosh, N., Gotway, M., Sil, A., et al. (2005). Bronchiectasis in persons with skin lesions resulting from arsenic in drinking water. *Epidemiology*, *16*(6), 760-765.

Hassan, M. M., Atkins, P., & Dunn, C. (2003). The spatial pattern of risk from arsenic poisoning: A Bangladesh case study. *Journal of Environmental Science and Health, Part A*, *38*(1), 1-24.

Hoque, M. A., Burgess, W. G., Shamsudduha, M., & Ahmed, K. M. (2011). Delineating low-arsenic groundwater environments in the Bengal Aquifer System, Bangladesh. *Applied Geochemistry*, *26*(4), 614-623.

Huang, L., Kulldorff, M., & Gregorio, D. (2007). A spatial scan statistic for survival data. *Biometrics*, *63*(1), 109-18.

Hunter, J. M. (1974). The challenge of medical geography. In J. M. Hunter (Ed.), *The Geography of Health and Disease* (pp. 1-31). Chapel Hill, NC: Unviversity of North Carolina at Chapel Hill Department of Geography.

ICDDRB. (2004). Arsenic contamination in Matlab, Bangladesh. *Health and Sciences Bulletin, 2*(4), 11-14.

ICDDRB. (2010). *Health and Demographic Surveillance System - Matlab* (Volume 42,.). Dhaka, Bangladesh: International Centre for Diarrhoeal Diseases Research, Bangladesh.

Jackman, S. (2011). pscl: Classes and methods for R developed in the Political Science Computational Laboratory, Stanford University. Stanford, CA.

Khan, M. M. H., Aklimunnessa, K., Kabir, M., & Mori, M. (2007). Determinants of drinking arsenic-contaminated tubewell water in Bangladesh. *Health Policy and Planning*, *22*(5), 335-343.

Klassen, A. C., Kulldorff, M., & Curriero, F. (2005). Geographical clustering of prostate cancer grade and stage at diagnosis, before and after adjustment for risk factors. *International Journal of Health Geographics*, *4*(1).

Kleinman, K. P., Abrams, A. M., Kulldorfff, M., & Platt, R. (2005). A model-adjusted space–time scan statistic with an application to syndromic surveillance. *Epidemiology and Infection*, *133*(3), 409-419.

Koenig, M. A., Bishai, D., & Khan, M. A. (2001). Health interventions and health equity: The example of measles vaccination in Bangladesh. *Population and Development Review*, *27*(2), 283-302.

Koenig, M. A., Rob, U., Khan, M. A., Chakraborty, J., & Fauveau, V. (1992). Contraceptive use in Matlab, Bangladesh in 1990: Levels, trends, and explanations. *Studies in Family Planning*, *23*(6), 352-364.

Kolenikov, S., & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth*, *55*(1), 128-165.

Kozul, C. D., Ely, K. H., Enelow, R. I., & Hamilton, J. W. (2009). Low-dose arsenic compromises the immune response to influenza A infection in vivo. *Environmental Health Perspectives*, *117*(9), 1441-1447.

Kozul, C. D., Hampton, T. H., Davey, J. C., Gosse, J. A., Nomikos, A. P., Eisenhauer, P. L., Weiss, D. J., et al. (2009). Chronic exposure to arsenic in the drinking water alters the expression of immune response genes in mouse lung. *Environmental Health Perspectives, 117*(7), 1108-1115.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics - Theory and Methods*, *26*(6), 1481-1496.

Kulldorff, M. (2010). SaTScan user guide for version 9.0.

Kulldorff, M., Feuer, E. J., Miller, B. A., & Freedman, L. S. (1997). Breast cancer clusters in the northeast United States: A geographic analysis. *American Journal of Epidemiology*, *146*(2), 161-170.

Kulldorff, M., & Information Management Services Inc. (2009). SaTScan v8.0: Software for the spatial and space-time scan statistics.

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, *34*(1), 1-14.

Lanata, C. F., & Black, R. E. (2008). Acute lower respiratory infections. In R. D. Semba, M. W. Bloem, & P. Piot (Eds.), *Nutrition and Health in Developing Countries* (pp. 179-214). Totowa, NJ: Humana Press.

Lemarie, A., Morzadec, C., Bourdonnay, E., Fardel, O., & Vernhet, L. (2006). Human macrophages constitute targets for immunotoxic inorganic arsenic. *The Journal of Immunology*, *177*, 3019-3027.

Loeb, M. B. (2003). Community-acquired pneumonia in older people: The need for a broader perspective. *Journal of the American Geriatrics Society*, *51*(4), 539-543.

Loeys, T., Moerkerke, B., De Smet, O., & Buysse, A. (2012). The analysis of zero-inflated count data: Beyond zero-inflated Poisson regression. *The British Journal of Mathematical and Statistical Psychology*, *65*(1), 163-180.

Margolis, P. A., Greenberg, R. A., Keyes, L. L., LaVange, L. M., Chapman, R. S., Denny, F. W., Bauman, K. E., et al. (1992). Lower respiratory illness in infants and low socioeconomic status. *American Journal of Public Health*, *82*(8), 1119-1126.

May, J. M. (1958). *The Ecology of Human Disease*. New York, NY: MD Publications, Inc.

Mayer, J. D. (2000). Geography, ecology and emerging infectious diseases. *Social Science & Medicine*, *50*(7-8), 937-952.

Mayer, J. D., & Meade, M. S. (1994). A reformed medical geography reconsidered. *The Professional Geographer*, *46*(1), 103-106.

Meade, M. S. (1977). Medical geography as human ecology: The dimension of population movement. *The Geographical Review*, *67*(4), 379-393.

Meade, M. S., & Earickson, R. J. (2000). *Medical Geography* (2nd ed.). New York, NY: The Guilford Press.

Milton, A. H., Asan, Z. H., Rahman, A., & Rahman, M. (2001). Chronic arsenic poisoning and respiratory effects in Bangladesh. *Journal of Occupational Health*, *43*, 136-140.

Mizgerd, J. P. (2006). Lung infection-a public health priority. *PLoS Medicine, 3*(2 e76), 0155-0158.

Mizgerd, J. P. (2008). Acute lower respiratory tract infection. *New England Journal of Medicine, 358*(7), 716-727.

Mosler, H.-J., Blöchliger, O. R., & Inauen, J. (2010). Personal, social, and situational factors influencing the consumption of drinking water from arsenic-safe deep tubewells in Bangladesh. *Journal of Environmental Management*, *91*(6), 1316-23.

National Research Council. (2001). *Arsenic in Drinking Water: 2001 Update*. Washington, DC: National Academy Press.

Nickson, R. T., Mcarthur, J. M., Ravenscroft, P., Burgess, W. G., & Ahmed, K. M. (2000). Mechanism of arsenic release to groundwater, Bangladesh and West Bengal. *Applied Geochemistry*, *15,* 403-413.

Olsen, C. E., Liguori, A. E., Zong, Y., Lantz, R. C., Burgess, J. L., & Boitano, S. (2008). Arsenic upregulates MMP-9 and inhibits wound repair in human airway epithelial cells. *American Journal of Physiology*, *295*(2), L293-302.

Olsen, S. F., Martuzzi, M., & Elliott, P. (1996). Cluster analysis and disease mapping--why, when, and how? A step by step guide. *BMJ*, *313*(7061), 863-866.

Oremland, R. S., & Stolz, J. F. (2003). The ecology of arsenic. *Science (New York, N.Y.), 300*(5621), 939-44.

Paul, B. K. (2004). Arsenic contamination awareness among the rural residents in Bangladesh. *Social Science & Medicine*, *59*(8), 1741-1755.

Paul, B. K., & De, S. (2000). Arsenic poisoning in Bangladesh: A geographic analysis. *Journal of the American Water Resources Association*, *36*(4), 799-809.

van der Poll, T., & Opal, S. M. (2009). Pathogenesis, treatment, and prevention of pneumococcal pneumonia. *Lancet*, *374*(9700), 1543-1556.

R Core Development Team. (2010). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rahman, A., Persson, L.-Å., Nermell, B., El Arifeen, S., Ekström, E.-C., Smith, A. H., & Vahter, M. (2010). Arsenic exposure and risk of spontaneous abortion, stillbirth, and infant mortality. *Epidemiology*, *21*(6), 797-804.

Rahman, A., Vahter, M., Ekström, E.-C., Rahman, M., Golam Mustafa, A. H. M., Wahed, M. A., Yunus, M., et al. (2007). Association of arsenic exposure during pregnancy with fetal loss and infant death: a cohort study in Bangladesh. *American Journal of Epidemiology*, *165*(12), 1389-1396.

Rahman, M., Vahter, M., Wahed, M. A., Sohel, N., Yunus, M., Streatfield, P. K., El Arifeen, S., et al. (2006). Prevalence of arsenic exposure and skin lesions. A population based survey in Matlab, Bangladesh. *Journal of Epidemiology and Community Health*, *60*(3), 242-248.

Raqib, R., Ahmed, S., Sultana, R., Wagatsuma, Y., Mondal, D., Hoque, A. M. W., Nermell, B., et al. (2009). Effects of in utero arsenic exposure on child immunity and morbidity in rural Bangladesh. *Toxicology Letters*, *185*(3), 197-202.

Ravenscroft, P., Burgess, W. G., Ahmed, K. M., Burren, M., & Perrin, J. (2005). Arsenic in groundwater of the Bengal Basin, Bangladesh: Distribution, field relations, and hydrogeological setting. *Hydrogeology Journal*, *13*(5-6), 727-751.

Ravenscroft, P., Howarth, R. J., & Mcarthur, J. M. (2006). Comment on "Limited Temporal Variability of Arsenic Concentrations in 20 Wells Monitored for 3 Years in Araihazar, Bangladesh." *Environmental Science & Technology*, *40*(5), 1716-1717.

Rose, C. E., Martin, S. W., Wannemuehler, K. A., & Plikaytis, B. D. (2006). On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. *Journal of Biopharmaceutical Statistics*, *16*(4), 463-481.

Rossman, T. G. (2007). Arsenic. In W. N. Rom & S. B. Markowitz (Eds.), *Environmental and Occupational Medicine* (4th ed., pp. 1006-1017). Philadelphia, PA: Lippincott Williams & Wilkins.

Rudan, I., Boschi-Pinto, C., Biloglav, Z., Mulholland, K., & Campbell, H. (2008). Epidemiology and etiology of childhood pneumonia. *Bulletin of the World Health Organization*, *86*(5), 408-416.

Savarimuthu, X., Hira-smith, M. M., Yuan, Y., von Ehrenstein, O. S., Das, S., Ghosh, N., Mazumder, D. N. G., et al. (2006). Seasonal variation of arsenic concentrations in tubewells in West Bengal, India. *Journal of Health, Population, and Nutrition*, *24*(3), 277-281.

Sazawal, S., & Black, R. E. (2003). Effect of pneumonia case management on mortality in neonates, infants, and preschool children: A meta-analysis of community-based trials. *Lancet Infectious Diseases*, *3*, 547-556.

Selgrade, M. K. (2007). Immunotoxicity: The risk is real. *Toxicological Sciences*, *100*(2), 328-332.

Sengupta, M. K., Mukherjee, A., Ahamed, S., Hossain, M. A., Das, B., Nayak, B., Chakraborti, D., et al. (2006). Comment on "Limited Temporal Variability of Arsenic Concentrations in 20 Wells Monitored for 3 Years in Araihazar, Bangladesh". *Environmental Science & Technology*, *40*(5), 1714-1715.

Simoes, E. A. F., Cherian, T., Chow, J., Shahid-Salles, S. A., Laxminarayan, R., & John, T. J. (2006). Acute respiratory infections in children. In D. T. Jamison, J. G. Breman, A. R. Measham, G. Alleyne, M. Claeson, D. B. Evans, P. Jha, et al. (Eds.), *Disease Control Priorities in Developing Countries* (2nd ed., pp. 483-497). Washington, DC: World Bank.

Smedley, P. L., & Kinniburgh, D. G. (2002). A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry*, *17*, 517-568.

Smith, A. H., Goycolea, M., Haque, R., & Biggs, M. L. (1998). Marked increase in bladder and lunger cancer mortality in a region of Northern Chile due to arsenic in drinking water. *American Journal of Epidemiology*, *147*(7), 660-669.

Smith, A. H., Lingas, E. O., & Rahman, M. (2000). Contamination of drinking-water by arsenic in Bangladesh: A public health emergency. *Bulletin of the World Health Organization, 78*(9), 1093-1103.

Smith, A. H., Marshall, G., Yuan, Y., Ferreccio, C., Liaw, J., von Ehrenstein, O., Steinmaus, C., et al. (2006). Increased mortality from lung cancer and bronchiectasis in young adults after exposure to arsenic in utero and in early childhood. *Environmental Health Perspectives, 114*(8), 1293-1296.

Smith, A. H., Marshall, G., Yuan, Y., Liaw, J., Ferreccio, C., & Steinmaus, C. (2011). Evidence from Chile that arsenic in drinking water may increase mortality from pulmonary tuberculosis. *American Journal of Epidemiology*, *173*(4), 414-420.

Smith, A. H., & Steinmaus, C. M. (2009). Health effects of arsenic and chromium in drinking water: Recent human findings. *Annual Reviews of Public Health*, *30*, 107-122.

Sohel, N., Persson, L. A., Rahman, M., Streatfield, P. K., Yunus, M., Ekström, E.-C., & Vahter, M. (2009). Arsenic in drinking water and adult mortality: A population-based cohort study in rural Bangladesh. *Epidemiology*, *20*(6), 824-830.

Sohel, N., Vahter, M., Ali, M., Rahman, M., Rahman, A., Streatfield, P. K., Kanaroglou, P. S., et al. (2010). Spatial patterns of fetal loss and infant death in an arsenic-affected area in Bangladesh. *International Journal of Health Geographics*, *9*(1).

Soto-Peña, G. A., Luna, A. L., Acosta-Saavedra, L., Conde, P., López-Carrillo, L., Cebrián, M. E., Bastida, M., et al. (2006). Assessment of lymphocyte subpopulations and cytokine secretion in children exposed to arsenic. *FASEB*, *20*(6), 779-781.

Spika, J. S., Munshi, M. H., Wojtyniak, B., Sack, D. A., Hossain, A., Rahman, M., & Saha, S. K. (1989). Acute lower respiratory infections: A major cause of death in children in Bangladesh. *Annal of Tropical Paediatrics*, *9*(1), 33-39.

Stewart, M. K., Fauveau, V., Parker, B., Chakraborty, J., & Kham, S. A. (1994). Acute respiratory infections in Matlab: Epidemiology, community perceptions and control strategies. In Vincent Fauveau (Ed.), *Matlab: Women, Children and Health* (pp. 187-204). Dhaka, Bangladesh: The International Centre for Diarrhoeal Disease Research, Bangladesh.

Vahter, M. (2002). Mechanisms of arsenic biotransformation. *Toxicology*, *182*, 211-217.

Vahter, M. (2009). Effects of arsenic on maternal and fetal health. *Annual Reviews of Nutrition*, *29*, 381-399.

Vahter, M., & Concha, G. (2001). Role of metabolism in arsenic toxicity. *Pharmacology & Toxicology*, *89*, 1-5.

WHO. (1991). *Technical Basis for the WHO Recommendations on the Management of Pneumonia in Children at First-level Health Facilities* (WHO/ARI/91.20). Geneva.

Wakefield, J., Kelsall, J. E., & Morris, S. E. (2001). Clustering, cluster detection, and spatial variation in risk. In P. Elliott, J. Wakefield, N. Best, & D. Briggs (Eds.), *Spatial Epidemiology*. Oxford: Oxford University Press.

Winans, B., Humble, M. C., & Lawrence, B. P. (2011). Environmental toxicants and the developing immune system: A missing link in the global battle against infectious disease? *Reproductive Toxicology*, *31*(3), 327-336.

Yunus, M., Sohel, N., Hore, S. K., & Rahman, M. (2011). Arsenic exposure and adverse health effects: A review of recent findings from arsenic and health studies in Matlab, Bangladesh. *The Kaohsiung Journal of Medical Sciences*, *27*(9), 371-6.

Zaman, K., Baqui, A. H., Yunus, M., Sack, R. B., Bateman, O. M., Chowdhury, H. R., & Black, R. E. (1997). Acute respiratory infections in children: A community-based longitudinal study in rural Bangladesh. *Journal of Tropical Pediatrics*, *43*, 133-137.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, *27*, 1-25.